

Capacity of Multiple Unicast in Wireless Networks: A Polymatroidal Approach

Sreeram Kannan* and Pramod Viswanath†

Coordinated Science Laboratory and Dept. of ECE
University of Illinois, Urbana-Champaign, IL 618101
Email: {kannan1, pramodv}@illinois.edu

Abstract

A classical result in undirected wireline networks is the near optimality of routing (flow) for multiple-unicast traffic (multiple sources communicating independent messages to multiple destinations): the min cut upper bound is within a logarithmic factor of the number of sources of the max flow. In this paper we “extend” the wireline result to the wireless context.

Our main result is the approximate optimality of a simple layering principle: *local physical-layer schemes combined with global routing*. We use the *reciprocity* of the wireless channel critically in this result. Our formal result is in the context of channel models for which “good” local schemes, that achieve the cut-set bound, exist (such as Gaussian MAC and broadcast channels, broadcast erasure networks, fast fading Gaussian networks).

Layered architectures, common in the engineering-design of wireless networks, can have near-optimal performance if the *locality* over which physical-layer schemes should operate is carefully designed. Feedback is shown to play a critical role in enabling the separation between the physical and the network layers. The key technical idea is the modeling of a wireless network by an undirected “polymatroidal” network, for which we establish a max-flow min-cut approximation theorem.

1 Introduction

Wireless networks are typically engineered using a layered approach: the physical layer deals with the channel noise, the medium access control layer deals with scheduling of users in the wireless context, the network layer handles routing of information and the transport layer deals with network congestion. While this design methodology has several engineering advantages that have led to the proliferation of wireless networks, a fundamental understanding of layering architectures is still lacking.

In this paper, we look explicitly for layered communication strategies that are near optimal for multiple unicast in general wireless networks. This would serve two (distinct) objectives:

- To obtain the approximate capacity of multiple unicast in wireless networks, and
- To establish a layered communication architecture that can guide engineering design.

*Supported in part by NSF grants CCF 1017430 and CNS 0721652.

†Supported in part by NSF grants CCF 1017430 and CNS 0721652.

1.1 Prior Work

Fundamental understanding of layering architectures has recently received plenty of attention from the networking community [55] [54], and scenarios have been identified under which a joint optimization of the transport and network layers naturally decompose into separate optimization problems, thus yielding a justification for the layered architecture. While there have been attempts to include certain aspects of the wireless medium into this framework [66], the understanding is far from complete. In this paper, we take a fundamental, information theoretic perspective, on if and when, the physical, medium access and network layers can be separately designed.

1.1.1 Capacity Results for Wireless Networks

Substantial progress has been made in the recent past in understanding the key aspects of the wireless medium (broadcast and superposition) from an information-theoretic view point. In particular, the capacity of MIMO broadcast channel has been resolved [92], approximate capacity of the 2-user interference channel has been established [21] and the approximate capacity of 2-user X -channels [59] [39], K -user interference channels with diversity [14], [65] [24, 64] has been obtained. While these results establish information theoretic understanding of several important (“physical layer”) channels, there is no conceptual guideline to fit the solutions for reliable communication for the channel in the context of a bigger network it could be a part of. In a different direction, there has been significant progress in understanding network-level capacity issues in the context of simple traffic models, starting from the breakthrough work [12], where the approximate capacity of single unicast is characterized, and later generalized to several scenarios: the approximate capacity of unicast in discrete memoryless networks is characterized in [51], a separation result between analog and digital components in relay networks is established in [10] and the approximate capacity of broadcasting in Gaussian networks is established in [2].

While unicast traffic in general Gaussian networks and multiple unicast traffic in single-hop Gaussian networks are reasonably well understood, the capacity of multiple unicast traffic in Gaussian wireless networks remains an open problem in multi-terminal information theory. In recent times, several research groups have made progress on this problem [80, 63, 40, 84, 11, 34], but the general problem still remains unsolved. Specific directions, with promise of success, involve simplifying the problem by considering specific traffic patterns such as 2-unicast [63, 80, 47, 90, 91, 41, 81, 32, 86]; another approach is to consider more specific network topologies, like for example, K sources communicating to K sinks via L fully-connected layers of K relays each [40, 70]. While these existing works attempt to compute the degrees-of-freedom (or approximate capacity) exactly for specific instances of the problem, we adopt a different viewpoint and focus our attention on obtaining general results for arbitrary networks (at the expense of obtaining potentially weaker approximation in specific instances).

In the context of multiple-unicast in large wireless networks, there has been significant progress in understanding *scaling laws* for geographical wireless networks; beginning with the seminal work in [33] and culminating in the hierarchical relaying scheme in [73] and a combination of the two [68, 69, 74] (with several critical works in between [95, 56, 7, 94, 27]). Despite its significant advantages, the performance guarantees are only in the context of certain specific wireless network models and, more importantly, the communication scheme is not a representation of a simple layered architecture for communication.

1.1.2 Information-theoretic Layering Architectures

Separation theorems form a basic tool in information theory: in his celebrated paper [79], Shannon showed that source coding (compression) and channel coding (communication) can be separated without loss of optimality. Following this, several separation (and non-separation) theorems have been proved in the multi-terminal context (see, for example, [18, 20]). The result most relevant to the current discussion is the sepa-

ration between network coding and channel coding proved in the pioneering work in [45]. There it is shown that, for a wireline network composed of *independent noisy channels*, a separation architecture composed of a physical layer that performs independent coding for each channel and a network layer which transports bits across the induced noiseless network, is optimal. This is a very interesting structural result that holds under arbitrary traffic models and is proved without the necessity to compute the capacity of the network. Thus the question of studying the capacity of wireline networks can be reduced to the question of studying the capacity of capacitated *graphs*.

For k -unicast in wireline graphs, a very interesting dichotomy is known in the theoretical computer science literature: for undirected graphs, the classical work of Leighton and Rao [50] shows that routing achieves the min-cut to within a $\mathcal{O}(\log k)$ factor and furthermore, there is a standing conjecture [52] [35] [43] [48] which claims that routing is infact optimal; whereas for general directed graphs, it has been shown recently [16] that there is no polynomial time algorithm that can approximate the value of min-cut to within a k^ϵ factor. Since max-flow can be computed in polynomial time, this result implies that there are networks for which flow cut gap is greater than k^ϵ . Furthermore, it has been proved recently [15] that computing the network coding region for directed graphs is equivalent to determining the entropy vector region, which is believed to be a very hard problem. Thus the string of positive results in the context of undirected (or bidirected) graphs and negative results in the context of directed graphs serve as an indicator that it may be easier to understand bidirected wireless networks.

A study of layering in directed wireless networks is initiated in [46]. The key idea is that of channel emulation, where a given channel is upper bounded by a wireline network with joint capacity constraints in such a way that the wireline network can emulate all possible behavior of the channel. This is a very strong condition, which ensures that the channel can be upper bounded by the wireline network *irrespective of the traffic pattern*. This program has been already accomplished for networks of 2-user MAC and broadcast channels, but appears to be very hard for general networks. In this paper, we instead work with only multiple-unicast traffic but go on to study layering in general bidirected networks.

1.1.3 Polymatroidal Networks

Polymatroidal networks and approximation theorems for these networks form a critical component in this paper. Polymatroidal networks are similar to edge capacitated networks, however, in addition to having edge capacity constraints, there are joint capacity constraints on the set of edges which meet at a vertex. Polymatroidal networks were introduced by Lawler and Martel [49] and Hassin [36] in the single-commodity setting and are closely related to the submodular flow model of Edmonds and Giles [23]; the well-known maxflow-mincut theorem holds in this more general setting (see Chapter 60 in [77] for a discussion). Multiple unicast in polymatroidal networks has not been studied till the recent work of [1], where an approximate max-flow min-cut result is established for undirected polymatroidal networks.

Recently, there have been several applications of network flows, and their generalizations such as flows in linking systems [78], to information flow in wireless networks [12, 9, 97, 31, 76, 42]. The polymatroidal structure of the multiple access channel capacity region was observed and exploited by Tse and Hanly [82]. Directed polymatroidal networks were utilized in the work of Vasudevan and Korada [85], where a separation architecture for a network of deterministic broadcast and MAC channels converts the network into a polymatroidal network and existing results for broadcasting in polymatroidal networks [25] are used to obtain capacity bounds.

1.2 Summary of Results

An important contribution of this paper is the following *meta-theorem*:

Meta Theorem: *If there is a “good” physical-layer scheme (which is approximately cut-set achieving and reciprocal) for a certain channel, then, for multiple-unicast in a network composed of such channels, a layered architecture is approximately cut-set achieving (to within a logarithmic factor in the number of messages).*

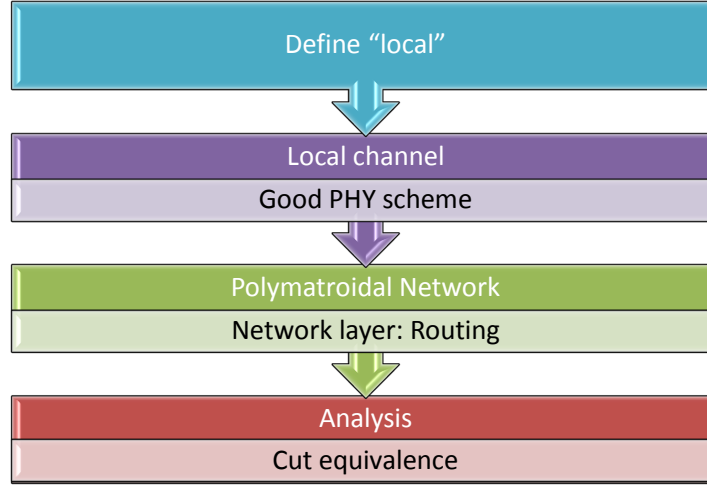


Figure 1: Layered Architecture

Fig. 1 depicts the layered architecture used in this paper. The following is a summary of the key ideas which are used in this paper to argue that a layered architecture is approximately capacity optimal:

- Model a wireless network as a bidirected network, by using the natural reciprocity of wireless networks.
- Utilize a good local “physical layer” scheme for each channel and identify the combinatorial structure of the rate region (typically submodularity).
- Show that local physical layer schemes convert a wireless network into a *bidirected polymatroidal network*. Thus the bidirected polymatroidal network can be viewed as a *graphical model* for wireless networks.
- Prove a Leighton-Rao type approximation result for bidirected polymatroidal network, which shows that routing is near optimal for k -unicast traffic.
- Argue that the layered architecture with local physical layer scheme + global routing achieves the cut-set approximately in the wireless network.
- We provide a technique by which “good” results for a given channel can be *lifted up* to good results for a general network comprised of those channels.

We justify the meta-theorem formally in the context of the following channel models:

1. Networks composed of Gaussian broadcast and MAC channels

2. Networks composed of broadcast erasure channels with feedback
3. Fast fading wireless networks
4. Degrees-of-freedom approximation for fixed wireless networks
5. Linear deterministic networks composed of MAC and broadcast channels
6. Networks composed of MIMO MAC and broadcast channels with delayed CSI feedback
7. Fast fading linear deterministic networks

For each of these networks, under a general k -unicast traffic model, the approximation factor on the rate is $\mathcal{O}(\log k)$ for the *entire rate region* in addition to the loss incurred due to the physical layer scheme, which is typically a power-scaling loss. Under more specific traffic models, such as the X -traffic model (where each of the J sources have messages to send to each of the K destinations) or a “group-communication” traffic model (where a subset S of nodes have messages to send to each other), we prove a constant approximation factor for the *sum-rate*, again in addition to a power scaling loss (the constant being 1, 2 or 4 depending on the specific channel and traffic model).

The material in this paper was presented in part in [4, 5].

1.3 Organization

The rest of this document is organized as follows:

- An overview of the layering approach is provided in Sec. 2. After defining the “locality” over which local physical layer schemes must be implemented, a list of desirable properties of local solutions are provided.
- In Sec. 3, polymatroidal networks, which form the back-bone of the layering architecture, are defined and results for multiple-unicast in polymatroidal networks are presented.
- In Sec. 4, “good” local physical layer solutions are described for various channel models. For some models, it is shown that existing schemes satisfy the desirable properties, whereas for other models, where existing schemes are insufficient, new ones are constructed.
- In Sec. 5, the local schemes are fitted into a global network context. Capacity theorems are proved for the various channel settings by connecting the wireless network problem formally to the polymatroidal network problem.

2 Layered Architecture

Engineering approaches to reliable network communication involve “layering”, a separation of the roles of physical (dealing with channel uncertainty), medium access (dealing with sharing the wireless medium) and networking (dealing with end-to-end the resulting “wireline” network communication). On the other hand, fundamental architectures are suggested by information theoretic study of large wireless networks (a major research direction in the past decade, with performance measured in a coarse scaling context). For instance, multihop routing [33] is a layered architecture, while hierarchical MIMO [73] (nearly scaling-law optimal in a geographically uniform context) is not. The information theoretic understanding of layering architectures has recently started receiving attention (see [45, 46]). Our approach is in similar lines as the approach in [85],

where a layered architecture for a network of deterministic broadcast and MAC channels is used to obtain capacity bounds.

In understanding the systematic design of layered architectures, it helps to look at the global wireless network as a collection of “local” wireless networks. The focus of this section is to introduce this view point; we propose that the notion of locality comes from both geographic (spatial) and temporal contexts. We see that certain combinatorial properties of the (physical layer) solutions to the local networks are *desirable*; these will help prove fundamental guarantees on the performance of the layered architecture in the global context.

2.1 Locality

A wireless network is a collection of *local* channels, if there are no interactions between the channels. Formally, a wireless network composed is defined as follows: Consider a graph $G = (V, E)$. For each $v \in V$, x_v, y_v denote the transmit and received symbol respectively. Let $C^+(v), C^-(v)$ denote the set of all channels in which v can transmit and receive from respectively, i.e., x_v can be written as $x_v = \{x_v^c\}_{c \in C^+(v)}$, and $y_v = \{y_v^c\}_{c \in C^-(v)}$. We will consider a wireless network with independent noise, where

$$\mathbb{P}(y_1, \dots, y_n | x_1, \dots, x_n) = \prod_v \mathbb{P}(y_v | \{x_u\}_{u: (u,v) \in E}), \quad (1)$$

and this description explicitly captures the relationship between the graph and the joint probability transition function.

Consider a set of channels $c \in \mathcal{C}$. A wireless network is said to be composed of channels $c \in \mathcal{C}$ if the probabilistic description of the network is of the form,

$$\mathbb{P}(y_1, \dots, y_n | x_1, \dots, x_n) = \prod_c \mathbb{P}(\{y_v^c\}_{v \in V^-(c)} | \{x_u^c\}_{u \in V^+(c)}) \quad (2)$$

$$= \prod_c \prod_{v \in V^-(c)} \mathbb{P}(y_v^c | \{x_u^c\}_{u \in V^+(c): (u,v) \in E}), \quad (3)$$

where $V^-(c) = \{v : c \in C^-(v)\}$ and $V^+(c) = \{v : c \in C^+(v)\}$. Each channel c is referred to as a component channel of the network.

A canonical scenario occurs when the wireless network is simply a collection of statistically independent noisy channels. Here each channel between a transmitter and a receiver is local. A more interesting example occurs in the case of a *frequency planned* wireless network, where each component of the wireless network operates in a specified frequency range. Here, the overall channel model can be decomposed as the product of channel models in each frequency range; the scale of locality corresponds to the scale of frequency reuse.

In general, such a geographic decomposition (via frequency planning) may not happen. Nevertheless, we can view the decomposition as occurring in *time* (indeed, this has been a popular method for analyzing general wireline / wireless networks [6, 12]). When we decompose across time, the local channel corresponds to the global one, as viewed over a specific single block of time. In this context, the layering architecture restricts the sophistication of physical layer (and medium access layer) strategies to be restricted to operate on a single layer in time, and at the end of each epoch, the information is *decoded* and re-encoded (using the networking layer) for the next local channel. The layering architecture thus enforces decoding of all information at each “hop” (in time); schemes, such as quantize-and-forward [12], which forward analog information do not fit the layered architecture model.

2.2 Desirable Properties of Local Solutions

A natural desirable property of any (physical layer) solution to a local channel is it be as optimal (from an information theoretic view point) as possible. In particular, we will be interested in how close the solution is to fundamental upper bounds given by the cutset bounds and certain natural combinatorial properties of the solution. For a network described by a probability transition matrix, the cut-set bound can be written as

follows. Given a cut $\Omega \subseteq V$, let D_Ω be the set of demands separated by the cut, i.e., $D_\Omega := \{k : s_k \in \Omega, t_k \in \Omega^c\}$. The cut-set bound bounds the sum of rates of sources in D_Ω and can be written as ¹

$$\sum_{k \in D_\Omega} R_k \leq \sup_{p_{x_1, \dots, x_n}} I(X_\Omega; Y_{\Omega^c} | X_{\Omega^c}). \quad (4)$$

Our focus on cut-set bounds as opposed to specialized outer bounds for specific wireless channels (such as the broadcast and interference channels) is motivated due to the following reasons.

- *Generality*: The cut-set bound [28] is an information theoretic outer bound on the achievable rate region and it can be written down for a general wireless network.
- *Decomposition*: The chain rule of mutual information allows the cut-set bound of a network to decompose into the cut-set bounds on local channels; thus solutions that come close to the cut-set bound at a local level have a potential to be layered and be still close to the cut-set bound at a global level. Formally, if we have a cut $\Omega \subseteq V$, the value of the cut is given by

$$I(X_\Omega; Y_{\Omega^c} | X_{\Omega^c}) = H(Y_{\Omega^c} | X_{\Omega^c}) - H(Y_{\Omega^c} | X_V) \quad (5)$$

$$= H(Y_{\Omega^c} | X_{\Omega^c}) - \sum_c H(Y_{\Omega^c}^c | X_V^c) \quad (6)$$

$$\leq \sum_c H(Y_{\Omega^c}^c | X_{\Omega^c}^c) - H(Y_{\Omega^c} | X_V) \quad (7)$$

$$\leq \sum_c H(Y_{\Omega^c}^c | X_{\Omega^c}^c) - H(Y_{\Omega^c} | X_V^c) \quad (8)$$

$$= \sum_c H(Y_{\Omega^c}^c | X_{\Omega^c}^c) - H(Y_{\Omega^c} | X_V^c) \quad (9)$$

$$= \sum_c I(X_\Omega^c; Y_{\Omega^c}^c | X_{\Omega^c}^c), \quad (10)$$

and thus the cut-set decomposes into sum of the cut-sets evaluated for each channel.

- *Structure*: Cut-set bounds have been well studied in the theoretical computer science literature and their combinatorial structure have been well understood. In fact, algorithms for approximately computing the cut-set bounds form an integral part of the theory of approximation algorithms.
- *Invariance under feedback*: The cut-set bound (evaluated under general joint distributions) is essentially a bound based on upper bounding the rate by the rate of a point-to-point channels and is, therefore, invariant to feedback.

Finally, *reciprocity* of the local channels (rate region reciprocity with the roles of transmitter and receiver reversed) will be paid attention to. The combinatorial structure imposed by the bidirected nature of each local channel will yield to efficient algorithms that are close to cuts.

2.3 Layering Methodology

Layering architectures stitch together the local solutions into a global solution:

1. The solution to a local channel allows for reliable digital communication at a local level.

¹While there is a stronger way of writing this bound, this weaker form of the bound will suffice for the purposes here.

2. Replacing each local channel by a set of (wireline) links leads to a network comprised of noiseless channels, with the rates on the various edges are *coupled* by the rate regions of the local solutions. Reciprocal local solutions ensure that the network obtained is *bidirected* (i.e., any edge between node a to node b has a corresponding edge between node b to a, with the two edges being involved in the same types of capacity region constraints).
3. Over the resulting wireline network, we might have to potentially employ network coding to re-encode the information between local channels. We utilize the combinatorial properties of the coupled rate constraints to study these new class of wireline networks. In particular, if the combinatorial structure governing the rate constraints is a specific form of a polytope known as a *polymatroid*, we obtain polymatroidal networks. Therefore, we study polymatroidal networks (which have local polymatroidal constraints on rate region) and prove that routing can achieve the cut-set bound to within a $\mathcal{O}(\log k)$ factor for the k -unicast problem, and also prove some better approximations for more specific communication problems.
4. Since the cut-set bound on a network of channels decomposes into a *sum* of cut-set bounds on the local channels, we can readily compare the performance of the layering architecture to a fundamental upper bound on the global network performance.

Whenever local solutions are close to the cut-set bounds for the corresponding local channels, we can establish the fundamental near-optimality of the layering architecture. We have accomplished this program for several canonical local wireless channels including broadcast erasure channels, Gaussian uplink and downlink channels, and interference channels with diversity (example: fast fading).

3 Polymatroidal Networks

Polymatroidal capacity networks and max-flow min-cut results for polymatroidal networks form the backbone of this paper. In this section, we give a brief introduction to polymatroidal networks. A detailed discussion of polymatroidal networks and the max-flow min-cut approximation result can be found in [1].

3.1 Polymatroids

A set function $f : 2^N \rightarrow \mathbb{R}$ over a finite ground set N is submodular iff $f(A) + f(B) \geq f(A \cap B) + f(A \cup B)$ for all $A, B \subseteq N$; equivalently $f(A \cup \{i\}) - f(A) \geq f(B \cup \{i\}) - f(B)$ for all $A \subseteq B$ and $i \notin A$. It is monotone if $f(A) \leq f(B)$ for all $A \subseteq B$. A polymatroid refers to the following set in \mathbb{R}^N :

$$\mathcal{P} = \left\{ (x_1, \dots, x_N) : \sum_{i \in S} x_i \leq f(S) \quad \forall S \subseteq [N] \right\}, \quad (11)$$

where $f(S)$ is a monotone submodular function with $f(\emptyset) = 0$. Thus a polymatroid is fully specified by specifying a monotone submodular function with $f(\emptyset) = 0$ (we will call such a function itself as a polymatroid). An example of a polymatroid is the following: given a set of N vectors v_1, \dots, v_N in \mathbb{R}^M , the function $f(S)$ defined as the rank of the matrix composed of $\{v_i\}_{i \in S}$ defines a polymatroid (we refer the reader to [71] for an introduction on polymatroids).

3.2 Definition of Polymatroidal Networks

A commonly studied wireline scenario is one where each edge is labeled by a capacity: this is the largest amount of information flowing on that edge. Here we are interested in a more general model which is able to

handle the additional constraints when edges meet at a node, similar in spirit to the broadcast and superposition constraints in wireless.

Consider a node v in a directed graph G and let $\delta_G^-(v)$ be the set of edges in to v and $\delta_G^+(v)$ be the set of edges out of v . In the standard model each edge (u, v) has a non-negative capacity $c(u, v)$ that is independent of other edges. In the polymatroidal network for each node v there are two associated submodular functions: ρ_v^- and ρ_v^+ which impose joint capacity constraints on the edges in $\delta_G^-(v)$ and $\delta_G^+(v)$ respectively. That is, for any set of edges $S \subseteq \delta_G^-(v)$, the total capacity available on the edges in S is constrained to be at most $\rho_v^-(S)$, similarly for $\delta_G^+(v)$. Note that an edge (u, v) is influenced by ρ_u^+ and ρ_v^- .

A cut $\Omega \subseteq V$ partitions the set of vertices into two sets Ω and Ω^c . Let $F(\Omega)$ denote the set of edges going from Ω to Ω^c (we will drop the dependence on Ω if it obvious from the context). The cost/capacity of F is complex to define in polymatroidal networks (note that in standard networks it is simply $\sum_{e \in F} c(e)$ where $c(e)$ is the capacity of e). To define the cost in polymatroidal networks, each edge (u, v) in F is first assigned to either u or v ; we say that an assignment of edges to nodes $g : F \rightarrow V$ is *valid* if it satisfies this restriction. A valid assignment partitions F into sets $\{g^{-1}(v) \mid v \in V\}$ where $g^{-1}(v)$ (the pre-image of v) is the set of edges in F assigned to v by g . Then we minimize over all assignments.

Definition 1. Cost of edge cut: *Given a directed polymatroidal network $G = (V, E)$ and a set of edges $F \subseteq E$, its cost denoted by $\nu(F)$ is $\min_{g: F \rightarrow V, g \text{ valid}} \sum_v (\rho_v^-(\delta^-(v) \cap g^{-1}(v)) + \rho_v^+(\delta^+(v) \cap g^{-1}(v)))$. In an undirected polymatroid network $\nu(F)$ is $\min_{g: F \rightarrow V, g \text{ valid}} \sum_v \rho_v(g^{-1}(v))$.*

A max-flow min-cut theorem for the unicast problem in directed polymatroidal networks is known in the literature [49]. This result has been extended to the broadcast traffic scenario in [25]. Our focus here is on multiple unicast traffic, for which flow-cut gaps are currently unknown. The best known result for multiple-unicast, even in the traditional wireline network with capacity constrained edges, is the approximate optimality of max flow (in terms of being close to the min cut); this result is available (in a strong sense) only for bidirected (or undirected) networks. Our goal is to obtain a similar result for bidirected polymatroidal networks.

We define a bidirected polymatroidal network as a directed polymatroidal network with the following properties.

- Every edge $e = (i, j)$ has a corresponding reverse edge $\tau(e) := (j, i)$.
- For any vertex v , the polymatroidal constraint $\rho_v^-(\cdot)$ on the incoming edges $\text{In}(v)$ is the same as the polymatroidal constraint $\rho_v^+(\cdot)$ on the outgoing edges $\text{Out}(v)$. More concretely,

$$\rho_v^-(E_v) = \rho_v^+(\tau(E_v)) \quad \forall E_v \subseteq \text{In}(v). \quad (12)$$

3.3 Main Result

The following theorem is proved in [1], which generalizes the results of [50] to the case of polymatroidal capacity networks:

Theorem 1. *For a bidirected polymatroidal network with k source-destination pairs, the ratio between the max-flow rate region and the min-cut rate region is $\mathcal{O}(\log k)$. The max-flow and an approximate min-cut can be calculated in polynomial time. Furthermore, this factor is tight in general, i.e., there are families of polymatroidal networks such that the flow-cut gap is $\Omega(\log k)$.*

Proof. (High level idea) The proof is done for undirected polymatroidal networks, whose capacity is within a factor 2 of bidirected capacity networks. The max-flow can be written down as a large linear program. While this program has an exponential number of constraints, a polynomial time algorithm can be obtained to solve this program using the polymatroidal structure. The dual of the linear program is related to the cut via an

integer relaxation. An important step in the proof is to simplify the constraint structure of the dual via the use of continuous extensions of submodular functions, in particular the Lovász extension [58]. The resulting program has a convex objective function but the constraint structure is much simpler. This convex program (which equals the dual of the flow) and the cut are related by an integer relaxation.

The work of Linial, London and Rabinovich [57] showed a fundamental (and *tight*) connection between flow-cut gaps in edge-capacitated undirected networks and low distortion embeddings² of finite metric spaces into ℓ_1^m . This connection effectively reduced the flow-cut gap question to investigating ℓ_1^m embeddings of finite metric spaces induced by graphs.

For the polymatroidal network problem, embeddings into a general ℓ_1 spaces are no longer sufficient, and we require a specific type of embedding called line embedding, where the nodes are embedded into a line in \mathbb{R} . These embeddings were implicit in the seminal work of Bourgain [13], first explored by Matousek and Rabinovich [62] [75] and later exploited in [26] to obtain flow-cut gaps in undirected node-capacitated graphs. These results imply that every finite metric space of n elements can be embedded by a contraction into a line preserving the average distortion of k points to within a factor of $\mathcal{O}(\log k)$. We use this result to connect the convex program written using the Lovász extension to the cut, hence proving a flow-cut gap of $\mathcal{O}(\log k)$.

The converse part, i.e., the existence of families of polymatroidal networks with a gap of $\Omega(\log k)$ follows from the existence of families of *edge capacitated* networks (which are special cases of polymatroidal networks) with a flow-cut gap of $\Omega(\log k)$. □

3.4 Special Traffic Scenarios

While in general, the factor of $\mathcal{O}(\log k)$ for flow-cut gaps is tight for multiple-unicast in bidirected polymatroidal networks, there may be special communication scenarios for multiple unicast when the factor can be improved. We present some instances here, where the flow cut gap is much better even for the more general case of *directed* polymatroidal networks.

3.4.1 Broadcast traffic:

Broadcast traffic is a special type of multiple unicast traffic where all the messages originate at a single source. Consider a directed polymatroidal network with a single source s having independent messages to K destination nodes t_1, \dots, t_K .

Lemma 1. [25] *For a directed polymatroidal network with broadcast-traffic pattern, the rate region of the max flow equals the rate region of the min-cut.*

3.4.2 Sum rate in directed X networks:

Consider J sources S_1, \dots, S_J and K destinations T_1, \dots, T_K , where each source has an independent message for each destination. The rate tuple is a JK length vector R_{jk} between each j and k . This communication problem is referred to commonly as the X -network problem.

Lemma 2. *For a directed polymatroidal network with X -traffic pattern, the sum-rate of max flow equals the sum-rate bound given by min-cut.*

Proof. Construct a super source S which talks to the J sources with infinite capacity links and a super sink T which is connected from each of the K sinks via infinite capacity links. The max-flow min-cut theorem for unicast between S and T in directed polymatroidal networks [49] implies the desired result. □

²An embedding of a metric space (V, d) (metric d on points V) into another metric space (V', d') is a mapping $f : V \rightarrow V'$. The distortion of f is the smallest $\alpha \geq 1$ such that $\frac{1}{\alpha} \cdot d(u, v) \leq d'(f(u), f(v)) \leq \alpha \cdot d(u, v)$ for all $u, v \in V$.

3.4.3 Sum rate for group communication in directed networks:

Consider a directed polymatroidal network with a specially marked group of nodes $S \subseteq V$. Each node s in S has an independent message for every other node in S . Thus it is a multiple unicast problem with $|S|(|S| - 1)$ messages. We refer to this traffic pattern as the group-communication traffic pattern. Suppose we are interested only in maximizing the sum-rate.

Lemma 3. [1] *For a directed polymatroidal network with a group-communication traffic pattern, the sum-rate of max-flow is greater than half the sum-rate bound given by min-cut.*

Proof. The proof of this theorem is non-trivial and requires a reduction from the directed polymatroidal network to the directed edge capacitated network using a combinatorial uncrossing argument. For a directed edge-capacitated network, this theorem is proved by Naor and Zosin [61]. For a detailed proof of this statement for polymatroidal networks, we refer the reader to [1]. \square

4 Local Physical Layer Schemes

In this section, good local physical layer schemes for several channel models will be discussed. For each of these channels, the goals will be to identify a physical layer scheme, quantify its rate region, understand its closeness to the cut-set bound and to examine its combinatorial structure. We will also analyze if the rate region remains (approximately) the same when the sources and destinations are exchanged and channels are reversed. In later sections, these properties will allow us to stitch together local physical layer schemes to get global schemes. The results in this section for various channel models are summarized in Table 1.

We will use the notation $\mathcal{R}_{\text{ach}}^{\text{ch}}$ and $\mathcal{R}_{\text{cut}}^{\text{ch}}$ to denote the achievable and the cut-set *rate regions* respectively, where ch denotes the channel of interest.

4.1 Linear Deterministic Broadcast and Multiple Access Channels

Consider a broadcast channel with d receivers of the form,

$$y_i = H_i x, \quad \forall i = 1, 2, \dots, d, \quad (13)$$

where y_i is the received vector at receiver i and x is the transmitted vector. The source intends to communicate independent messages to each of its destinations. For a subset K of $\{1, 2, \dots, d\}$, let H_K denote the matrix with $H_i, i \in K$ stacked up alongside one another. The capacity region of this broadcast channel [60] is given by

$$\mathcal{C} = \{(R_1, \dots, R_d) : \sum_{i \in K} R_i \leq \text{Rank}(H_K) \quad \forall K \subseteq [d]\}. \quad (14)$$

This capacity region is also equal to the cut-set bound, which is a polymatroid (see [71]).

Let us consider a “reciprocal” multiple access channel in which there are d transmitters and one receiver,

$$y = \sum_i H_i^T x_i, \quad (15)$$

and all the transmitters have an independent message to transmit to the single destination. The capacity region of a general MAC channel is known (see, for example, Chapter 14 in [17]) and for the linear deterministic channel, is given again by the rate region in (14). We observe that the capacity of the broadcast channel and the reciprocal MAC channel are the same and are equal to their cut-set bound.

Thus for a linear deterministic broadcast and MAC channel, the rate region is exactly polymatroidal, equal to the cut, and is reciprocal.

Characteristic / Channel	Closeness-to-Cut	Combinatorial Structure	Reciprocity
Linear Deterministic MAC / BC	Exact	Polymatroidal	Exact
Gaussian MAC / BC	Approximate	Polymatroidal	Approximate
Erasure Broadcast	Far	Polytope	Far
Erasure Broadcast (Feedback)	Approximate	Polymatroidal	Approximate
Fading MAC / BC with delayed CSI	Approximate	Polymatroidal	Approximate
Fading linear deterministic	Approximate	Polymatroidal	Approximate
Fading X-Channel	Approximate	Polymatroidal	Approximate
Fixed X-Channel	Approximate in DOF	Polymatroidal DOF region	Approximate in DOF

Table 1: Canonical wireless channels, as viewed via three lenses.

4.2 Gaussian Broadcast and Multiple Access Channels

Let us first consider a multiple access channel, defined by

$$y = \sum_i h_i x_i + z, \quad (16)$$

where the transmitted vector x_i is constrained by a power constraint P at each of the d nodes, y is the received vector and z denotes the noise, which is of unit power. Let the rate region achievable on this multiple access channel be denoted by $\mathcal{R}_{\text{ach}}^{\text{MAC}}(P)$. This region is known to be polymatroidal [82]. This rate region equals the cut-set bound evaluated under product distributions [17], i.e.,

$$\mathcal{R}_{\text{ach}}^{\text{MAC}}(P) = \mathcal{R}_{\text{cut,product}}^{\text{MAC}}(P). \quad (17)$$

Let the cut-set bound evaluated under general distributions be given by $\mathcal{R}_{\text{cut,general}}^{\text{MAC}}(P)$. We can easily verify the relation,

$$\mathcal{R}_{\text{ach}}^{\text{MAC}}(P) \supseteq \mathcal{R}_{\text{cut,general}}^{\text{MAC}}\left(\frac{P}{d}\right). \quad (18)$$

Next, let us consider a “reciprocal” or “dual” broadcast channel, given by,

$$y_i = H_i x + z_i, \quad \forall i = 1, 2, \dots, d, \quad (19)$$

where the transmitted vector x is constrained by a power constraint dP , y_i is the received vector and z_i denotes the noise at each receiver, which is of unit power. Let us call the rate region of the broadcast channel

as $\mathcal{R}^{\text{BC}}(P)$. This rate region has been fully characterized, but is not equal to the cut-set bound and is not polymatroidal (see Chapter 6 in [83] for a discussion).

The rate region of this broadcast channel with sum power constraint Pd contains the rate region of the multiple access channel with *individual constraint* P at each node [87, 88], i.e.,

$$\mathcal{R}_{\text{ach}}^{\text{BC}}(P) \supseteq \mathcal{R}_{\text{ach}}^{\text{MAC}}(P). \quad (20)$$

For purpose of symmetry, we can choose to operate the broadcast channel at the rate region specified by $\mathcal{R}_{\text{ach}}^{\text{MAC}}(P)$, i.e., let us set

$$\mathcal{R}_{\text{ach}}^{\text{BC}}(P) = \mathcal{R}_{\text{ach}}^{\text{MAC}}(P). \quad (21)$$

This will also ensure that the rate region of the achievable scheme $\mathcal{R}_{\text{ach}}^{\text{BC}}(P)$ is polymatroidal. Thus, in our achievable strategy, the rate region of a multiple access and that of the dual broadcast channel are equal and given by a polymatroidal region.

Let the cutset bound for the broadcast channel be specified as $\mathcal{R}_{\text{cut,general}}^{\text{BC}}(P)$. Since there is only one input variable, the cut-set bound under product distribution and general distribution are the same.

Lemma 4. *The achievable region of the MAC channel compares with the cutset bound under general distributions as follows,*

$$\mathcal{R}_{\text{ach}}^{\text{MAC}}(P) \supseteq \mathcal{R}_{\text{cut,general}}^{\text{MAC}}\left(\frac{P}{d}\right). \quad (22)$$

For the broadcast channel, we have the relation,

$$\mathcal{R}_{\text{ach}}^{\text{BC}}(P) \supseteq \mathcal{R}_{\text{cut,general}}^{\text{BC}}\left(\frac{P}{d}\right). \quad (23)$$

Proof. The proof is deferred to Appendix A. □

Thus for a Gaussian broadcast and MAC channel, the rate region is approximately polymatroidal, close to the cut, and is approximately reciprocal.

4.3 Broadcast Erasure Channels

Consider a network comprised of broadcast erasure channels. For a broadcast erasure channel with d receivers, the channel model can be written as,

$$y_i = e_i x, \quad \forall i = 1, 2, \dots, d, \quad (24)$$

where e_i is a binary random variable which when 0 represents that at receiver i , the packet got erased. If e_i are all independent, then the broadcast erasure channel is said to be an independent erasure broadcast channel. Such a channel is specified by erasure probabilities $\epsilon_i, i = 1, 2, \dots, d$, where

$$\epsilon_i = \Pr\{e_i = 0\}. \quad (25)$$

For the purpose of simplicity in this paper we will consider only broadcast erasure channels that are independent and symmetric, which implies that there is only one parameter ϵ and $\epsilon_i = \epsilon \quad \forall i$.

Lemma 5. *For an erasure broadcast channel without feedback, the cut-set bound can be as large as L times the achievable rate region.*

Proof. See Appendix B. □

This result implies that for broadcast erasure channels (without feedback), there are no good local schemes that can achieve close to the cut.

4.3.1 Broadcast Erasure Channels with Feedback

Since there are no good local schemes for the broadcast channel, we suggest that the scale of the locality be enlarged to include the presence of feedback links in the physical model to get better local schemes.

The capacity of the erasure broadcast channel with ACK feedback (the receiver acknowledges whether it received the packet) was studied by [30] for the two user channel, and later extended to the more general case independently by [89] and [29]. The schemes are based on network coding and interference alignment and demonstrate that the following rate region is achievable.

Lemma 6. [89, 29] *The following rate region is achievable for the erasure broadcast channel with ACK feedback,*

$$\mathcal{R}_{ach,fb} = \left\{ (R_1, \dots, R_D) \mid \sum_{i=1,2,\dots,d} \frac{R_{\pi(i)}}{1 - \epsilon^i} \leq 1 \quad \forall \pi \right\}. \quad (26)$$

Here π is a permutation of the set $(1, \dots, d)$. Note that this region is not polymatroidal. However, it is close to the min cut rate region (which is itself polymatroidal) as seen below:

Lemma 7.

$$\mathcal{R}_{ach,fb} \supseteq \frac{\mathcal{R}_{cut}}{\mathcal{O}(\log d)}. \quad (27)$$

Proof. See Appendix C. □

The reciprocal nature of wireless channels from which the broadcast erasure channel is constructed naturally suggests a way of providing feedback links of *commensurate* strength. Formally: a channel is said to have *commensurate feedback* if there are feedback links from the various receiving nodes to the transmitters with the same rate region as the cut-set bound for the forward channel. In Appendix D, we look at one possible way of obtaining feedback links of commensurate strength as the forward links.

Thus for a broadcast erasure channel with feedback, the rate region is approximately polymatroidal, close to the cut, and is approximately reciprocal.

4.4 MIMO Broadcast and MAC Channel with Delayed Feedback

We will consider MIMO broadcast and MAC channels, similar to Sec. 4.2, but with the difference that the channel states are i.i.d. over antennas and time. If the CSI is instantaneously available at the transmitter and the receiver, then the methods and results of Sec. 4.2 continue to hold. However, the problem becomes interesting when global channel state information is no longer available. In some settings, the channel may change so fast that by the time, the channel state feedback reaches the transmitter, the channel state has changed significantly. This feature is essentially present in the erasure channel if we think of erasure as being a channel state, which is unknown to the transmitter originally but the presence of the ACK feedback delivers this channel state to the transmitter with a delay. Recent work by Ali and Tse [8] showed the surprising result that even in Gaussian networks, delayed CSI can be very beneficial as compared to the absence of CSIT [96]. This will form the basis of our investigation of these channels.

We assume that each broadcast and MAC channel gets feedback from its receivers about the channel state, however, the channel changes before the feedback arrives, precluding the use of feedback to predict the future state of the channel. We will resort to a degree-of-freedom characterization for this problem. Let d_i denote the achievable degrees of freedom for the i -th message, i.e.,

$$d_i := \lim_{\text{SNR} \rightarrow \infty} \frac{R_i(\text{SNR})}{\log \text{SNR}}, \quad (28)$$

where R_i is the achievable rate for user i . The achievable DOF region is denoted by \mathcal{D}_{ach} , and the region given by the cutset bound is denoted by \mathcal{D}_{cut} .

Such a multiple access case is easy to deal with because even without channel state information, the cut-set bound can be achieved,

$$\mathcal{D}_{\text{ach}}^{\text{MAC}} = \mathcal{D}_{\text{cut}}^{\text{MAC}}. \quad (29)$$

However for the corresponding broadcast channel, in the absence of CSIT, we can only achieve rates that are far from the cut-set. In [8], it is shown that the presence of channel state feedback, even when delayed, can significantly alter the situation, as was the case with broadcast erasure channels:

Lemma 8. [8] *For a fading MISO broadcast channel with a source with L transmit antennas and L single transmit antenna receivers, the following DOF region can be achieved with the help of delayed CSIT:*

$$\mathcal{D}_{\text{ach}}^{\text{BC}} = \left\{ \sum_{i=1}^L D_i \leq l \frac{K}{\sum_{i=1}^L \frac{1}{i}} \right\}. \quad (30)$$

Our first result is to obtain an approximation for the rate region of MIMO broadcast channels with delayed feedback, formally stated in the following lemma.

Lemma 9. *For a fading MIMO broadcast channel with l transmit antennas and d users with user i having m_i antennas, and delayed CSI feedback, the following DOF region can be achieved:*

$$\mathcal{D}_{\text{ach}}^{\text{BC}} \supseteq \frac{\mathcal{D}_{\text{cut}}^{\text{BC}}}{\mathcal{O}(\log p)}, \quad (31)$$

where $p := \min(l, \sum_i m_i)$ is the minimum of the number of transmit and receive antennas in the system.

Thus for a fading MIMO broadcast channel, the rate region is approximately polymatroidal, close to the cut, and is approximately reciprocal.

4.5 Fading X-channels

Consider a L -user X -channel, there are L sources s_1, \dots, s_L and L destinations t_1, \dots, t_L with *each* source having message to send to each destination, thus there are L^2 messages in total. The connectivity graph between the nodes is a bi-partite graph with E being the set of edges. We will abbreviate edge (s_i, t_j) as (i, j) since the meaning is clear from the context.

Note that an interference channel is a special case of this channel. Capacity achieving schemes even for the L -user interference channel are not known in the general setting. In [14], the authors show that each user in a L -user interference channel can achieve half their point-to-point degrees-of-freedom (DOF) if the channel is fast-fading, using a mechanism called interference alignment, which was initially proposed for the 2-user X -channel in [59]. This result has since been generalized in several directions; most notably, in [65], it is shown that using an “ergodic interference alignment” scheme, each user can get half her rate at *all* SNR, and in [64], it is shown that the DOF result can be proved even under *fixed channel coefficients* using a scheme termed “real interference alignment”. These results have been unified into a single framework in [93]. It has also been shown [67] that ergodic interference alignment can be used to achieve linear capacity scaling in dense interference networks.

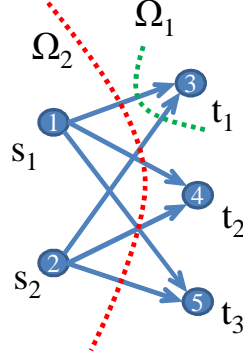


Figure 2: Cuts in X -Channel

4.5.1 Cuts in an X -channel

Consider an example X -channel with two sources and three destinations, shown in Fig. 2. Two cuts are marked in the figure. The light (green) cut Ω_1 separates only destination 1 from the two sources, thus providing a bound: $R_{11} + R_{21} \leq C_{\Omega_1}$, whereas dark (red) cut Ω_2 separates all sources from all destinations and therefore provides a bound on $R_{11} + R_{12} + R_{13} + R_{21} + R_{22} + R_{23} \leq C_{\Omega_2}$. Of these two bounds, the first bound corresponds to that of a polymatroidal constraint whereas the second bound does not correspond to a polymatroidal constraint (since in a polymatroidal network, only edges that meet at a node have a joint constraint).

For a general X -channel, these two types of cuts will be present, and we can classify them as

1. Cuts that separate a single node from the rest of the nodes (referred to as cuts of the polymatroidal form), and
2. Cuts that separate multiple sources from multiple destinations.

We would like to show an achievable scheme that not only achieves the cut-set bound approximately, but also the rate region of the achievable scheme satisfies a polymatroidal constraint. Therefore, for an X -channel, we will have to show that only cuts of the polymatroidal form (separating one node from the rest) play a dominant role. This is a key challenge that we address in this section.

4.5.2 Channel Model

The channel model can therefore be written as,

$$y_i(t) = \sum_{j \in \text{In}(i)} h_{ij}(t)x_j(t) + z_i(t) \quad \forall t = 1, 2, \dots, T, \quad (32)$$

where $x_i(t)$, $y_i(t)$, $z_i(t)$ are the transmitted vector, received vector and noise vector at time t , $\text{In}(i)$ represents the set of neighbors of node i who have an incoming edge to i and fading coefficient $h_{ij}(t)$ is associated with edge $(i, j) \in E$ at time t . The noise vector is assumed to have unit variance and is independent at each node. There is a power constraint of P per node.

We will make the following assumptions about the fading distribution:

- Fading coefficients are assumed to be i.i.d. over edges and over time.

- The fading coefficient will be assumed to be symmetric, i.e., if h_{ij} is a discrete random variable,

$$\Pr\{h_{ij} = a\} = \Pr\{h_{ij} = -a\}, \quad \forall a \quad (33)$$

otherwise, if the random variable is absolutely continuous, the pdf $p(\cdot)$ must satisfy

$$p(h_{ij} = a) = p(h_{ij} = -a) \quad \forall a. \quad (34)$$

- The fading distribution is assumed to satisfy:

$$a := e^{-\mathbb{E}(\log |h|^2)} < \infty. \quad (35)$$

One example of a fading distribution that satisfies these assumptions is when $h_{ij}(t)$ is i.i.d. across nodes and time with a complex gaussian distribution, for which $a = 1.723$ [72].

We will use the shorthand $C(P)$ to denote the ergodic capacity of a fading channel with power constraint P , and the fading coefficient h of unit variance,

$$C(P) := \mathbb{E}_h \left[\frac{1}{2} \log(1 + |h|^2 P) \right]. \quad (36)$$

4.5.3 Scheme for the K -user interference channel

First, we consider the case of k -user interference channel, where there are messages only from s_i to t_i for each $i = 1, 2, \dots, k$. In the ergodic set-up, the following result is known:

Lemma 10. [65] *For an k -user ergodic interference channel, where the direct links are non-zero, the following rate tuple is achievable:*

$$R_i = \frac{1}{2} C(2P), \quad \forall i = 1, 2, \dots, k. \quad (37)$$

4.5.4 Scheme for the X -channel

We generalize this physical layer scheme to the X -channel with L sources and M destinations, demonstrating not only that the cut-set bound is approximately achievable, but also that only cut of the polymatroidal form are relevant.

Theorem 2. *For an L -source, M -sink ergodic X -channel, the following rate region is achievable,*

$$\mathcal{R}_X^{ach} = \left\{ (R_{ij}) \mid \begin{array}{l} \sum_{j:(i,j) \in E} R_{ij} \leq \frac{1}{2} C(2P) \quad \forall i \in S \\ \sum_{i:(i,j) \in E} R_{ij} \leq \frac{1}{2} C(2P) \quad \forall j \in T \end{array} \right\}. \quad (38)$$

and furthermore, if d is the maximum degree of any node,

$$\mathcal{R}_{ach}^X(P) \supseteq \frac{\mathcal{R}_{cut}^X(\frac{2P}{ad})}{2}, \quad (39)$$

where $a := e^{-\mathbb{E}(\log |h|^2)}$.

Proof. Let us write R_{ij} for the rate of communication between s_i and t_j . We use the following achievable strategy:

- Let us construct a bipartite graph between the source vertices and sink vertices, and edges given by E .

- A matching in a bipartite graph is a choice of edges such that each node is present in at most one edge. In our case, a matching can be thought of representing a choice of at most one destination for each source. Choose a matching M on the bipartite graph, let π be the corresponding permutation. The characteristic vector of a bipartite matching is given by the vector $(x_{ij}) : x_{ij} = 1$, if $(i, j) \in M$, otherwise $x_{ij} = 0$.
- Consider the interference channel from s_1, \dots, s_L to $d_{\pi(1)}, \dots, d_{\pi(L)}$. For the $s_i, t_{\pi(i)}$ pairs that are connected, we can achieve a rate of $\frac{1}{2}C(2P)$ using the strategy of Lemma 10.
- This implies that a rate given by $\frac{1}{2}C(2P)$ times the characteristic vector of the bipartite matching is achievable.
- Now, we can achieve any convex combination of the rates given by matchings on the graph. This is given by the following polytope, called the matching polytope,

$$\mathcal{M} = \text{conv}\{(x_{ij})_M | M \text{ a matching}\}. \quad (40)$$

- By a theorem in bi-partite graph matchings [77], this matching polytope can be alternately described as:

$$\mathcal{P} = \left\{ (x_{ij}) | x_{ij} \geq 0 \quad \forall i, j, \quad \sum_{j:(i,j) \in E} x_{ij} \leq 1 \quad \forall i, \quad \sum_{i:(i,j) \in E} x_{ij} \leq 1 \quad \forall j \right\}. \quad (41)$$

- Therefore the achievable rate region is given by (51).
- The cut-set bound $\mathcal{R}_X^{\text{cut}}$ implies the following, which are only a subset of the cuts (the cuts which separate one node from all the others) :

$$\left\{ (R_{ij}) | \begin{array}{l} \sum_{j:(i,j) \in E} R_{ij} \leq \mathbb{E} \log(1 + \sum_{j:(i,j) \in E} h_{ij}^2 P) \quad \forall i \in S \\ \sum_{i:(i,j) \in E} R_{ij} \leq \mathbb{E} \log(1 + \sum_{i:(i,j) \in E} h_{ij}^2 P) \quad \forall j \in T \end{array} \right\}.$$

- Now, due to the concavity of the logarithm,

$$\mathbb{E} \log(1 + \sum_{j:(i,j) \in E} h_{ij}^2) \leq \log(1 + \sum_{j:(i,j) \in E} \mathbb{E} h_{ij}^2 P) \quad (42)$$

$$\leq \log(1 + dP) \quad (43)$$

$$\leq \mathbb{E} \log(1 + adP|h|^2) = C(adP) \quad (44)$$

where the last step follows because of the convexity of the function $f(x) = \log(1 + ce^x)$, i.e., applying Jensen inequality for the aforementioned convex function, we get,

$$\mathbb{E}(\log(1 + c|h|^2)) = \mathbb{E}\{\log(1 + ce^{\log|h|^2})\} \quad (45)$$

$$\geq \log(1 + ce^{\mathbb{E}(\log|h|^2)}) \quad (46)$$

$$= \log(1 + ca^{-1}), \quad (47)$$

where $a := e^{-\mathbb{E}(\log|h|^2)}$.

Thus the cut-set bound implies the following inequalities,

$$\left\{ (R_{ij}) | \begin{array}{l} \sum_{j:(i,j) \in E} R_{ij} \leq C(adP) \quad \forall i \in S \\ \sum_{i:(i,j) \in E} R_{ij} \leq C(adP) \quad \forall j \in T \end{array} \right\}.$$

- Therefore we get the result that

$$\mathcal{R}_{\text{ach}}^X(P) \supseteq \frac{\mathcal{R}_{\text{cut}}^X(\frac{2P}{ad})}{2}. \quad (48)$$

□

Thus for a fading X channel, the rate region is exactly polymatroidal, approximately close to the cut, and is exactly reciprocal (since the description of the rate region remains the same even the channel is reversed).

4.6 Fixed X -channels

Consider a L -user X -channel with fixed channel coefficients drawn from a continuous distribution. We will obtain a degrees-of-freedom characterization of this X -channel (which holds almost surely). The channel model can therefore be written as,

$$y_i(t) = \sum_{j \in \text{In}(i)} h_{ij} x_j(t) + z_i(t) \quad \forall t = 1, 2, \dots, T, \quad (49)$$

where $x_i(t)$, $y_i(t)$, $z_i(t)$ are the transmitted vector, received vector and noise vector at time t , $\text{In}(i)$ represents the set of neighbors of node i who have an incoming edge to i and channel coefficient h_{ij} associated with edge $(i, j) \in E$ is drawn from a continuous distribution which has a probability density function, i.e., the probability measure is absolutely continuous with respect to the Borel measure. The noise vector is assumed to have unit variance and is independent at each node. There is a power constraint of P per node.

First, we consider the case of L -user interference channel, where there are messages only from s_i to t_i for each i . The following result characterizes the degrees of freedom of the L -user interference channel:

Lemma 11. [64] *For a k -user interference channel with channel coefficients drawn from a continuous distribution, if the direct links are non-zero, the following DOF tuple is achievable almost surely:*

$$D_i = \frac{1}{2}, \quad \forall i = 1, 2, \dots, k. \quad (50)$$

We can generalize this interference channel scheme to the X -channel with L sources and M destinations, using the same method as in Theorem 2.

Theorem 3. *For an L -source, M -sink fixed X -channel, the following DOF region is achievable almost surely,*

$$\mathcal{D}_X^{\text{ach}} = \left\{ (D_{ij}) \mid \begin{array}{ll} \sum_{j: (i,j) \in E} D_{ij} \leq \frac{1}{2} & \forall i \in S \\ \sum_{i: (i,j) \in E} D_{ij} \leq \frac{1}{2} & \forall j \in T \end{array} \right\}. \quad (51)$$

and furthermore,

$$\mathcal{D}_{\text{ach}}^X \supseteq \frac{\mathcal{D}_{\text{cut}}^X}{2} \quad \text{a.s.} \quad (52)$$

Proof. The proof proceeds in a manner quite similar to that of Theorem 2; the only difference is that we use Lemma 11 instead of Lemma 10. Also, since we are dealing with DOF, which is an SNR-scaling characterization, the (constant) power scaling factor is not relevant. □

Thus for a fixed X channel, the achievable DOF region is exactly polymatroidal, approximately close to the cut, and is exactly reciprocal (since the description of the achievable DOF region remains the same even the channel is reversed).

4.7 Fading Linear Deterministic Channels

In Sec. 4.1, we considered linear deterministic networks which had only broadcast and MAC components. In this section we will consider a general linear deterministic network with fading. The communication network is represented by a directed graph $G = (V, E)$. If $(i, j) \in E \iff (j, i) \in E$, we call the network bidirected. The edges (i, j) that are present have fading matrix $H_{ij}(t)$ on them. Each fading coefficient in each matrix is distributed i.i.d. fading over edges and time. The fading distribution for each non-zero coefficient is assumed to be uniform over the finite field (except zero) \mathbb{F}_q . The proof can be extended to the case where the fading takes value 0 but there will be penalty factor of $\frac{q-1}{q}$.

For a linear deterministic interference network with fading, there is a scheme based on ergodic interference alignment that achieves half the point-to-point rate for each user:

Lemma 12. [65] *For an k -user fading linear deterministic interference channel, with direct links being non-zero, the following rate type is achievable:*

$$R_i = \frac{1}{2} \log_2(q), \quad \forall i = 1, 2, \dots, k. \quad (53)$$

We can use this scheme to create a scheme for the X -channel, the rate region of this scheme is quantified in the following theorem.

Theorem 4. *For an L -source, M -destination deterministic ergodic X -channel, the following rate region is achievable,*

$$\mathcal{R}_{ach}^X \supseteq \frac{\mathcal{R}_{cut}^X}{2}, \quad (54)$$

furthermore, only cuts that separate one node from the rest are sufficient.

Proof. The proof for this case is very similar to the proof of Theorem 2, except that the term $\frac{1}{2}C(2P)$ is now replaced by $\frac{1}{2} \log_2(q)$. Also, there are no power scaling losses in the case of a linear deterministic network. \square

Thus for a linear deterministic X channel with fading, good local schemes exist; furthermore the schemes have a rate region which can be described by polymatroidal constraints.

5 Approximate Capacity Results for Wireless Networks

In this section, we will present approximate capacity results for wireless networks with *multiple unicast traffic* for several channel models. Our achievable schemes will use a layering architecture and we will use the cut-set bound as the outer bound. For each channel model, we will have the description of the wireless network as a graph $G = (V, E)$. There are k designated source nodes s_1, \dots, s_k , which have independent messages and k corresponding destination nodes t_1, \dots, t_k : t_i wants to decode the message of s_i with vanishingly small probability of error, for every $i = 1, \dots, k$. Let \mathcal{R}_{ach}^{ch} denote the rate region comprising of the rate tuples achievable and let \mathcal{R}_{cut}^{ch} denote the rate region corresponding to the cut-set bound for the channel model ch . If a power constraint P is present, we will describe these regions as a function of P , i.e., $\mathcal{R}_{ach}^{ch}(P)$ and $\mathcal{R}_{cut}^{ch}(P)$.

5.1 Gaussian networks with MAC and Broadcast components

We consider Gaussian networks in which there are only broadcast and MAC components, i.e., there is no interference channel component. This is equivalent to the assertion that each edge is involved in either a

superposition constraint or in a broadcast constraint but not in both. In practice, such a network can be realized by using a partial frequency reuse scheme, where the total bandwidth is divided into different chunks, which are assigned to users in such a way that interference component is avoided.

Gaussian networks with MAC and broadcast components alone have been previously considered in [46], where it has been shown that for a network composed of 2-user MAC and broadcast channels, a separation architecture involving local coding is approximately optimal. Since this general structural result holds for all possible traffic models, it may be necessary in general to use network coding [6] [44] at the network layer. In contrast, in this paper, we will assume multiple unicast traffic and utilize the *reciprocity* of the wireless network to show that the cut-set bound can be approximately achieved using a separation scheme along with *routing*.

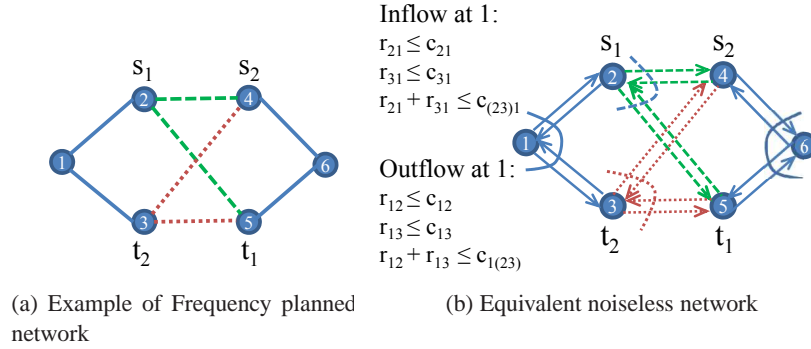


Figure 3: A Network composed of broadcast and MAC components and its Polymatroidal Equivalent

Network Model: The communication network is represented by an undirected graph $G = (V, E)$, and an edge coloring $\psi : E \rightarrow C$, where C is the set of colors. Each node v has a set of colors $C(v) \subseteq C$ on which it transmits and receives. Each color can be thought of as an orthogonal resource (for example, a frequency band), and therefore the broadcast and superposition constraints for the wireless channel apply only *within* a given color. We will assume the nodes are equipped with full duplex radios on each of these resources. For simplicity of notation, we will assume that there is a one-to-one correspondence between colors and channels, i.e., each channel operates on a distinct color; so c stands for a unique color and a unique channel.

The channel model can therefore be written as,

$$y_i^c = \sum_{j \in \mathcal{N}_c(i)} h_{ij}^c x_j^c + z_i^c \quad \forall c \in C(i), \quad (55)$$

where x_i^c, y_i^c, z_i^c are the transmitted vector, received vector and noise vector on color c , $h_{ij}^c = h_{ji}^c$ is the channel coefficient between node i and node j on color c and $\mathcal{N}_c(i)$ represents the set of neighbors of node i who are operating on color c and $d_c(i) = |\mathcal{N}_c(i)|$ be the degree of node i in color c . Let $d = \max_{c,v} d_c(v)$ be the maximum degree of any node in a given color; therefore, d is the maximum number of users on any component broadcast or multiple access channel. Each node has a power constraint P per edge. Therefore node v has power constraint $Pd_c(v)$ for transmitting on color c . By the very definition, this network has a reciprocal MAC channel for every broadcast channel and vice versa. Let $V(c) = \{v : c \in C(v)\}$ be the set of nodes that use the color c . An example of a wireless network along with its equivalent noiseless network are shown in Fig. 3.

Theorem 5. *For the k -unicast problem in Gaussian network composed of broadcast and multiple access channels, a simple separation strategy can achieve a rate,*

$$\mathcal{R}_{ach}^g(P) \supseteq \frac{\mathcal{R}_{cut}^g(\frac{P}{d})}{\mathcal{O}(\log k)}. \quad (56)$$

This means that the min cut, scaled down in power by a factor d and in rate by a factor $\mathcal{O}(\log k)$, can be achieved. For the unicast scenario ($k = 1$), we can show using a similar proof that

$$\mathcal{R}_{ach}^g(P) \supseteq \mathcal{R}_{cut}^g\left(\frac{P}{d}\right). \quad (57)$$

This result is similar to that obtained by [12], except that here it is obtained for the special case of networks composed of broadcast and multiple access channels. The scheme in [12] requires a global physical layer scheme (the “quantize and map” strategy), while for the special case of networks here we show that a simple separation strategy suffices.

5.1.1 Coding Scheme: Proof of Theorem 5

The coding scheme is a separation-based (layered) strategy: each component broadcast or multiple access channel is coded for independently creating bit-pipes on which information is routed globally. The achievable scheme used for the MAC and broadcast channel are discussed in detail in Sec 4.2. The achievable rate region for the MAC and Broadcast channels described there are polymatroidal and therefore each multiple access or broadcast channel with d users can be replaced by a set of d bit-pipes whose rates are jointly constrained by the corresponding polymatroidal constraints. Thus we get a polymatroidal network by using this layered strategy; this polymatroidal network is described as follows: for each node v in the original graph, there are several vertices v_c , one for each color $c \in C(v)$. There is an edge between u_c and v_c if $h_{uv}^c \neq 0$, the polymatroidal constraints are given by

$$\rho_{v_c}^-(F_v) = \log(1 + \sum_{u:(u,v) \in F_v} |h_{uv}^c|^2 P) \quad \forall F_v \subseteq \delta^-(v_c) \quad (58)$$

$$\rho_{v_c}^+(F_v) = \log(1 + \sum_{u:(u,v) \in F_v} |h_{vu}^c|^2 P) \quad \forall F_v \subseteq \delta^+(v_c), \quad (59)$$

and the polymatroidal network is bidirected due to the fact that $h_{uv}^c = h_{vu}^c$ and the reciprocity in the rate regions of the MAC and BC channel. Further, there are edges between v_c and $v_{c'}$ of infinite capacity, since these correspond to the same node v in the original graph.

Let us call the cut-set bound region on this induced polymatroidal network as $\mathcal{R}_{cut}^{\text{poly}}(P)$, and let us call the rate region for the flow-based achievable scheme as $\mathcal{R}_{ach}^{\text{poly}}(P)$. Then we have from Theorem 1 that

$$\mathcal{R}_{flow}^{\text{poly}}(P) \supseteq \frac{\mathcal{R}_{cut}^{\text{poly}}(P)}{\mathcal{O} \log k}. \quad (60)$$

As an example, the bidirected polymatroidal network induced for the example of Fig. 3a is shown in Fig. 3b. The submodular constraints are explicitly written down only for node 1, but similar constraints apply at nodes 2, 3 and 6. In this figure, c_{21} , c_{31} and $c_{(23)1}$ represent constraints on the rate of communication from node 2 to 1, the rate of communication from node 3 to 1 and the sum rate from nodes 2 and 3 to 1 respectively.

It is now sufficient to compare the cuts on the polymatroidal and the Gaussian network.

Lemma 13.

$$\mathcal{R}_{cut}^g(P) \subseteq \mathcal{R}_{cut}^{\text{poly}}(dP). \quad (61)$$

Proof. Given a cut F_Ω in the polymatroidal network, we will show that there is a corresponding cut in the Gaussian network, whose value is within a power scaling factor d of the polymatroidal cut. The value of the cut in the polymatroidal network is $\nu(F_\Omega) = \sum_c \nu(F_\Omega^c)$, i.e., the polymatroidal cut breaks up into the sum of the cuts of various colors. The value of cut in a given color $\nu(F_\Omega^c)$ corresponds to a certain polymatroidal constraint in a given MAC or broadcast channel. We need to show that there is a similar cut in the Gaussian network whose value is within a power scaling factor.

As shown in Lemma 4, we have that for each of these channels, the Gaussian cut and the polymatroid representing the achievable scheme are within a power scaling factor of d , and therefore

$$\mathcal{R}_{\text{cut}}^g(P) \subseteq \mathcal{R}_{\text{cut}}^{\text{poly}}(dP), \quad (62)$$

since both polymatroidal and Gaussian cuts decompose into sum of individual cuts. \square

The achievable rate using the separation strategy is given using (60) as

$$\mathcal{R}_{\text{ach}}^{\text{poly}}(P) \supseteq \frac{\mathcal{R}_{\text{cut}}^{\text{poly}}(P)}{\mathcal{O}(\log k)} \supseteq \frac{\mathcal{R}_{\text{cut}}^g(\frac{P}{d})}{\mathcal{O}(\log k)}. \quad (63)$$

This completes the proof of Theorem 5.

5.1.2 Special Traffic Scenarios

We now present results for *directed* networks with MAC and broadcast components under the special traffic patterns presented in Sec. 3.4.

Theorem 6. *For a directed Gaussian network composed of broadcast and multiple access channels, a simple separation strategy can achieve a rate,*

$$\mathcal{R}_{\text{ach}}(P) \supseteq \mathcal{R}_{\text{cut}}\left(\frac{P}{d}\right) \quad \text{for BC Traffic}, \quad (64)$$

$$R_{\text{ach}}^{\text{sum}}(P) \supseteq R_{\text{cut}}^{\text{sum}}\left(\frac{P}{d}\right) \quad \text{for X Traffic}, \quad (65)$$

$$R_{\text{ach}}^{\text{sum}}(P) \supseteq \frac{R_{\text{cut}}^{\text{sum}}\left(\frac{P}{d}\right)}{2} \quad \text{for group-communication traffic}. \quad (66)$$

Proof. The proofs are based on the polymatroidal results Lemma 1, Lemma 2 and Lemma 3. The proof is very similar to the proof of Theorem 5 and are therefore omitted. \square

5.2 Broadcast Erasure Networks with Commensurate Feedback

Broadcast erasure networks, in which there are broadcast but no superposition constraints, serve as high level models for communication in wireless networks. Unicast in broadcast erasure networks is well understood, for which it has been shown [19] that min-cut is achievable using a global linear network coding scheme (in [53], it is shown that knowledge of erasure locations is not necessary at the destination). It has also been shown that a separation scheme in which each broadcast erasure channel is coded for locally to create noiseless links does not perform very well. This is due to the fact that for each broadcast erasure channel the capacity region is far away from the min-cut region. However, as shown in Sec. 4.3, by utilizing ACK feedback, the capacity region is enlarged to become closer to the min-cut region. Therefore we consider a network of broadcast erasure channel, where each channel has a feedback mechanism.

Consider a network which is composed of broadcast erasure channels, with an appropriate mechanism for feedback built into the network. In particular, we look for feedback that is commensurate; formally: a

channel is said to have *commensurate feedback* if there are feedback links from the various receiving nodes to the transmitters with the same rate region as the cut-set bound for the forward channel. The reciprocal nature of wireless channels from which the broadcast erasure channel is constructed naturally suggests a way of providing feedback links of commensurate strength. In Appendix D, we look at one possible way of obtaining feedback links of commensurate strength as the forward links.

From now on, we will assume that each erasure broadcast channel has commensurate feedback, without cognizance to how this particular rate region was obtained. For simplicity, we will assume that all broadcast erasure channels are symmetric and independent, i.e., each broadcast erasure channel has erasure independent probability ϵ .

Theorem 7. *For the k unicast problem in a network of erasure broadcast channels with commensurate feedback, a simple separation strategy can achieve a rate*

$$\mathcal{R}_{ach}^{erasure} \supseteq \frac{\mathcal{R}_{cut}^{erasure}}{\mathcal{O}(\log k)\mathcal{O}(\log d_{\max})}, \quad (67)$$

where d_{\max} is the maximum number of users in any broadcast erasure channel.

Proof. Consider the following separation strategy: even though the feedback links have a rate region given by R_{cut}^{BC} , we will restrict them to use the rate region $R_{ach,fb}^{BC}$ in order to preserve symmetry. The feedback links are used for two distinct purposes:

- To provide Ack / Nack feedback. This feedback has an overhead of 1-bit per packet which we treat as negligible. This assumption makes sense especially when packet lengths are large.
- To route flows on the reverse direction, since the Ack feedback overhead is assumed to be small, this will essentially occupy the whole capacity. We establish a bidirected network by using the feedback links for routing.

Since we have the Ack / Nack feedback for each erasure broadcast channel, we can use the scheme of Lemma 6 to obtain the rate region $R_{ach,fb}^{BEC}$ with feedback. This induces a bidirected polymatroidal network specified in the following manner, in which we can use flows to achieve a rate region R_{ach}^{poly} . By Theorem 1, we have that

$$R_{ach}^{poly} \supseteq \frac{\mathcal{R}_{cut}^{poly}}{\mathcal{O}(\log k)}. \quad (68)$$

By Lemma 7, we have $\mathcal{R}_{ach,fb}^{BEC} \supseteq \frac{\mathcal{R}_{cut}^{BEC}}{\mathcal{O}(\log d_{\max})}$. Further, since cuts in the polymatroidal and the original network decompose into cuts for each channel, any cut-set in the polymatroidal network induced by the achievable scheme has a counter-part cut-set in the erasure network within a factor of $\mathcal{O}(\log d_{\max})$:

$$\mathcal{R}_{cut}^{poly} = \frac{\mathcal{R}_{cut}^{erasure}}{\mathcal{O}(\log d_{\max})}. \quad (69)$$

Now (68) and (69) together imply,

$$\mathcal{R}_{ach}^{poly} \supseteq \frac{\mathcal{R}_{cut}^{erasure}}{\mathcal{O}(\log k)\mathcal{O}(\log d_{\max})}, \quad (70)$$

which proves the desired result. \square

5.2.1 Special Traffic Scenarios

We now present results for *directed* networks with broadcast erasure channels under the special traffic patterns presented in Sec. 3.4. Since the networks are directed, reciprocity is not needed; however we will continue to assume that the broadcast erasure channel has ACK feedback.

Theorem 8. *For a directed network composed of broadcast erasure channels with ACK feedback, a simple separation strategy can achieve a rate,*

$$\mathcal{R}_{ach} \supseteq \frac{\mathcal{R}_{cut}}{\log(d_{\max} + 1)} \quad \text{for BC Traffic,} \quad (71)$$

$$R_{ach}^{sum} \supseteq \frac{R_{cut}^{sum}}{\log(d_{\max} + 1)} \quad \text{for X Traffic,} \quad (72)$$

$$R_{ach}^{sum} \supseteq \frac{R_{cut}^{sum}}{2 \log(d_{\max} + 1)} \quad \text{for group-communication traffic,} \quad (73)$$

where d_{\max} is the maximum degree of the broadcast channel.

Proof. The proofs are similar to the proof of Theorem 7 and are therefore omitted. \square

5.3 Gaussian Fast Fading Network

We will now consider a general Gaussian network where broadcast and superposition can simultaneously occur, i.e., the network can contain interference components. Such a network is clearly more general than the K -user interference channel, and even this channel has not been well understood in the most general case (the tutorial [37] provides an excellent summary of the current understanding on this channel). However, in the presence of fast fading, the problem gets symmetrized considerably [65], and there is a reasonable understanding of this problem. Therefore we will resort to the fast fading model in this section. We will further assume that the fading distribution satisfies the assumptions in Sec. 4.5.

While most of the existing literature is on single hop interference channels, multi-hop interference networks have been the focus of more recent work. In particular, it has been shown in [38] that the degrees of freedom of such fully-connected layered networks can be achieved using a non-separation scheme called opportunistic interference alignment. It has also been shown [32] that a separation architecture does not even achieve the DOF for simple networks, for example, the network with 2 sources, 2 relays and 2 destinations. Our results offer a contrast: if we look to achieve within an approximation factor of cut-set then a simple separation strategy suffices, for all SNR.

For examples of networks considered here, we refer the reader to Fig. 4 for a non-layered example or the one in Fig. 5 for a layered example.

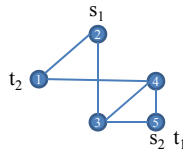


Figure 4: A Multiple-Unicast Wireless Network

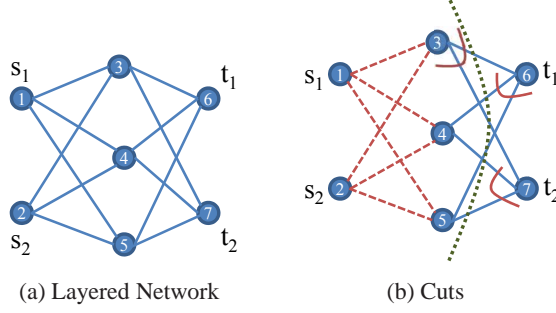


Figure 5: Layered Network and Cuts

Theorem 9. *For a bidirected ergodic wireless network with k source destination pairs, the rate region given by*

$$\mathcal{R}^g(P) \supseteq \frac{\mathcal{R}_{\text{cut}}^g\left(\frac{P}{bd^3}\right)}{\mathcal{O}(\log k)}, \quad (74)$$

is achievable using a separation strategy, where d is the maximum degree of any node and $b := \frac{e^{-\mathbb{E}(\log |h|^2)}}{2}$ is a constant depending on the fading distribution. If the fading is complex Gaussian, $b \approx 0.86$.

Proof. We will now show a scheme by which we can convert the bidirected ergodic wireless network into a bidirected polymatroidal network. We will view one snapshot of transmission in the network as a transmission in a bipartite graph, from the set of nodes $V^l = V$ on the left side to the set of nodes V^r on the right side. Node u on the left is connected to node v on the right with channel coefficient h_{uv} from the original network. The nodes have infinite memory, so each node is connected to itself with infinite capacity. We view the obtained network as a single hop network and use a scheme for the single hop (described in Sec. 4.5). The achievable rates for this single hop network are given by polymatroidal constraints, and therefore we obtain a polymatroidal network represented by a bi-partite graph with $2V$ nodes. This process can now be *reverted*, i.e., we can go from a layered representation with $2V$ nodes to a non-layered representation with V nodes. Thus we obtain a polymatroidal network with V nodes. There are edges E similar to the original graph, however the capacities are constrained according to the polymatroidal constraint:

$$\sum_{u \in \text{In}v} R_{uv} \leq \frac{1}{2}C(2P) \quad \forall v \in V \quad (75)$$

$$\sum_{u \in \text{Out}v} R_{vu} \leq \frac{1}{2}C(2P) \quad \forall u \in V. \quad (76)$$

This is now a bidirected polymatroidal network since the polymatroidal constraint is symmetric for the incoming and the outgoing edges at any given node. Now we perform routing over this bidirected polymatroidal network. Now the flow and cut on this network are related by:

$$\mathcal{R}_{\text{ach}}^g = \mathcal{R}_{\text{ach}}^{\text{poly}} \supseteq \frac{\mathcal{R}_{\text{cut}}^{\text{poly}}}{\mathcal{O}(\log k)}. \quad (77)$$

If we can relate the cut-set bound on the polymatroidal network and the cut-set bound on the original Gaussian network, we can get the desired result. This is done in the following lemma:

Lemma 14.

$$\mathcal{R}_{cut}^{poly}(P) \supseteq \frac{1}{2} \mathcal{R}_{cut}^{original} \left(\frac{P}{bd^3} \right). \quad (78)$$

Proof. See Appendix F □

Now, (77) and (78) implies that

$$\mathcal{R}_{ach}^{original} \supseteq \frac{\mathcal{R}_{cut}^{original} \left(\frac{P}{bd^3} \right)}{\mathcal{O}(\log k)}, \quad (79)$$

which completes the proof of the theorem. □

5.3.1 Multi-Antenna Nodes

We can prove a result similar to the one in Theorem 9 even when there are multiple antennas at the nodes.

Theorem 10. *For a bidirected ergodic wireless network with multiple antenna nodes, the rate region for k -unicast is given by*

$$\mathcal{R}(P) \supseteq \frac{\mathcal{R}_{cut} \left(\frac{P}{bd^3} \right)}{\mathcal{O}(\log k)}, \quad (80)$$

is achievable using a separation strategy, where d is the maximum number of antennas that can communicate with any given antenna and $b := \frac{e^{-\mathbb{E}(\log |h|^2)}}{2}$ is a constant depending on the fading distribution.

Proof. The proof of the theorem is very similar to that of Theorem 9 except that when there are multiple antennas, each antenna is treated as a separate node with infinite capacity wireline links between antennas of the same node. This scheme can be shown to achieve the cut-set of this new network to within the approximation factor. We observe that any finite cut in this new network will partition the antennas in such a way that all antennas corresponding to the same original node will lie on the same side of the partition, since otherwise the value of the cut will become infinite. Therefore the cut-set in the new network and the original network have the same value and this completes the proof of the theorem. □

5.3.2 Special Traffic Scenarios

We now present results for *directed* fast fading networks under the special traffic patterns presented in Sec. 3.4. Since the network is directed, reciprocity will not be necessary to prove this result.

Theorem 11. *For a directed fast fading Gaussian network with multiple antenna nodes, a simple separation strategy can achieve a rate,*

$$\mathcal{R}_{ach}(P) \supseteq \frac{\mathcal{R}_{cut} \left\{ \frac{P}{bd^3} \right\}}{2} \quad \text{for BC Traffic,} \quad (81)$$

$$R_{ach}^{sum}(P) \supseteq \frac{R_{cut}^{sum} \left\{ \frac{P}{bd^3} \right\}}{2} \quad \text{for X Traffic,} \quad (82)$$

$$R_{ach}^{sum}(P) \supseteq \frac{R_{cut}^{sum} \left\{ \frac{P}{bd^3} \right\}}{4} \quad \text{for group-communication traffic,} \quad (83)$$

where d is the maximum number of antennas that can communicate with any given antenna and $b := \frac{e^{-\mathbb{E}(\log |h|^2)}}{2}$ is a constant depending on the fading distribution.

Proof. The proofs are similar to the proof of Theorem 9 and are therefore omitted. Here, the factor loss of 2 is present for the BC scenario due the factor 2 loss in the local physical layer scheme (see Theorem 2). □

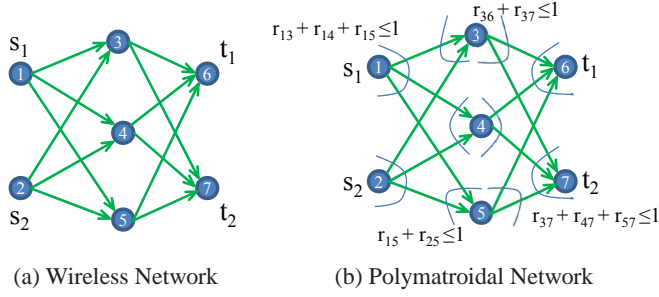


Figure 6: Directed Layered F.C. Networks

5.3.3 Special Channel Model: Directed Layered Network

In this section, we consider directed fully-connected (f.c.) layered networks. These are layered networks, which have connectivity between adjacent layers only in the forward direction, i.e., links are always between a node in V_i to a node in V_{i+1} . Further for a fully-connected network, we assume that $(u, v) \in \mathcal{E} \quad \forall u \in V_i, v \in V_{i+1}$. Consider, for example, the network in Fig. 6a.

Theorem 12. *For a directed fully-connected layered ergodic wireless network with k -distinct sources in the first layer having messages to k distinct sinks in the last layer, the rate region given by*

$$\mathcal{R}(P) \supseteq \frac{\mathcal{R}_{\text{cut}} \left\{ \frac{P}{bd^3} \right\}}{2}, \quad (84)$$

is achievable using a separation strategy, where d is the maximum degree of any node and $b := \frac{e^{-\mathbb{E}(\log |h|^2)}}{2}$.

Proof. The achievable scheme is again a local Phy scheme, i.e., for each hop in the layered network we use the strategy for X-channels described in Sec. 4.5. Now we get a special type of directed polymatroidal network for which we can show that max-flow equals min-cut.

Theorem 13. *For a k -unicast problem in a polymatroidal directed layered network in which all the node constraints are of the form*

$$\sum_{u \in \text{In}\{v\}} R_{uv} \leq 1, \quad \forall v, \quad (85)$$

$$\sum_{v \in \text{Out}\{u\}} R_{uv} \leq 1, \quad \forall u, \quad (86)$$

the rate region given by max-flow equals the min-cut for the k -unicast problem.

Proof. See Appendix G. □

Now we can use Lemma 14 to get:

$$\mathcal{R}_{\text{ach}}^{\text{original}}(P) = \mathcal{R}_{\text{ach}}^{\text{poly}}(P) = \mathcal{R}_{\text{cut}}^{\text{poly}}(P) \supseteq \frac{1}{2} \mathcal{R}_{\text{cut}}^{\text{original}} \left\{ \frac{P}{bd^3} \right\}, \quad (87)$$

which implies the theorem. □

5.4 Fixed Gaussian Network

Consider a general Gaussian network where broadcast and superposition can occur simultaneously, and where the channel coefficients are fixed but drawn from a continuous distribution. We studied a simple single-hop version of such a network in Sec. 4.6. We will now show that the local scheme can be placed in a network context to get an almost-sure DOF characterization.

Theorem 14. *For a bidirected fixed multi-antenna Gaussian network with k source destination pairs, the DOF given by*

$$\mathcal{D}_{ach} \supseteq \frac{\mathcal{D}_{cut}}{\mathcal{O}(\log k)}, \quad (88)$$

is achievable almost surely using a separation strategy.

Proof. The proof proceeds analogously to Theorem 10, but instead of using the local scheme in Theorem 2, the scheme of Theorem 3 is used and therefore a DOF characterization is obtained. \square

5.4.1 Special Traffic Scenarios

We now present results for *directed* Gaussian networks under the special traffic patterns presented in Sec. 3.4. Since the network is directed, reciprocity will not be necessary to prove this result.

Theorem 15. *For a directed Gaussian network with each channel coefficient chosen from a continuous distribution, a simple separation strategy can achieve a DOF region,*

$$\mathcal{D}_{ach} \supseteq \frac{\mathcal{D}_{cut}}{2} \quad \text{for BC Traffic,} \quad (89)$$

$$D_{ach}^{sum} \supseteq \frac{D_{cut}^{sum}}{2} \quad \text{for X Traffic,} \quad (90)$$

$$D_{ach}^{sum} \supseteq \frac{D_{cut}^{sum}}{4} \quad \text{for group-communication traffic.} \quad (91)$$

5.5 Linear deterministic networks with MAC and Broadcast components

A linear deterministic network composed of MAC and broadcast components is defined in the same way as the Gaussian network composed of MAC and broadcast components. The key difference is that the transmissions are over a finite field \mathbb{F}_q , there is no noise, and the channels between node i and j of color c are matrices in general: H_{ij}^c . The main result for linear deterministic networks with MAC and broadcast components is stated in the following theorem:

Theorem 16. *For the k -unicast problem in linear deterministic network composed of broadcast and multiple access channels, the layered architecture can achieve a rate,*

$$\mathcal{R}_{ach}^{ld} \supseteq \frac{\mathcal{R}_{cut}^{ld}}{\mathcal{O}(\log k)}. \quad (92)$$

Proof. The proof of this theorem proceeds in a similar way as that of Theorem 5, the key difference being the fact that for linear deterministic networks, the cut-set bounds under product form and general distributions are the same, and therefore, there is no power scaling factor. \square

5.5.1 Special Traffic Scenarios

We now present results for *directed* linear deterministic networks with MAC and broadcast components under the special traffic patterns presented in Sec. 3.4.

Theorem 17. *For a directed linear deterministic network composed of broadcast and multiple access channels, a simple separation strategy can achieve a rate,*

$$\mathcal{R}_{ach} \supseteq \mathcal{R}_{cut} \text{ for BC Traffic,} \quad (93)$$

$$R_{ach}^{sum} \supseteq R_{cut}^{sum} \text{ for X Traffic,} \quad (94)$$

$$R_{ach}^{sum} \supseteq \frac{R_{cut}^{sum}}{2} \text{ for group-communication traffic.} \quad (95)$$

Proof. The proofs are similar to the proof of Theorem 16 and are therefore omitted. \square

5.6 Networks of Fast Fading MAC and Broadcast Channels with delayed CSIT

We will now networks composed of fast fading MAC and Broadcast channels where the channel states are i.i.d. over antennas and time, and the CSI is available at the transmitting nodes after a delay. Schemes for each of these local channels was studied in Sec. 4.4 and a degree of freedom characterization was obtained. Now, we will try to obtain the degree of freedom region for multiple unicast over a network of such channels.

Our main result for such networks is the following.

Theorem 18. *For the k unicast problem in Gaussian network composed of fading broadcast and multiple access channels with delayed feedback, a simple separation strategy can achieve a DOF,*

$$\mathcal{D}_{ach} \supseteq \frac{\mathcal{D}_{cut}}{\mathcal{O}(\log k)\mathcal{O}(\log p_{\max})}, \quad (96)$$

where

$$p_{\max} = \max_{BC \text{ channels}} p(BC \text{ channel}), \quad (97)$$

and $p(BC \text{ channel})$ is given by the minimum of number of transmit antennas and the total number of received antennas in the broadcast channel.

Proof. The coding scheme is again a separation-based strategy: each component broadcast or multiple access channel with delayed feedback is coded for independently creating bit-pipes on which information is routed globally. The physical layer technique is the scheme for MIMO broadcast and MAC channels with delayed CSIT proposed in Sec. 4.4. The proof is very similar to that in Sec. 5.2.

For the fading MIMO broadcast channel with delayed feedback, we can achieve the following DOF region \mathcal{D}_{ach}^{BC} with feedback. For the fading MIMO multiple access channel, we can achieve the region given by \mathcal{D}_{ach}^{MAC} , however we will restrict ourselves to achieve the smaller rate region \mathcal{D}_{ach}^{BC} for the purpose of symmetry. This induces a bidirected polymatroidal network, in which we can use flows to achieve a rate region R_{ach}^{poly} . By Theorem 1, we have

$$R_{ach}^{poly} \supseteq \frac{\mathcal{R}_{cut}^{poly}}{\mathcal{O}(\log k)}. \quad (98)$$

By Lemma 9, we have $\mathcal{D}_{ach}^{BC} \supseteq \frac{\mathcal{D}_{cut}^{BC}}{\mathcal{O}(\log p_{\max})}$ and by choice, $\mathcal{D}_{ach}^{MAC} = \frac{\mathcal{D}_{cut}^{MAC}}{\mathcal{O}(\log p_{\max})}$. Further, since cuts in the polymatroidal and the original fading network decompose into cuts for each channel, any cut-set in the

polymatroidal network induced by the achievable scheme has a counter-part cut-set in the erasure network within a factor of $\mathcal{O}(\log p_{\max})$:

$$\mathcal{R}_{\text{cut}}^{\text{poly}} = \frac{\mathcal{D}_{\text{cut}}^{\text{fading}}}{\mathcal{O}(\log p_{\max})}. \quad (99)$$

Also, we can achieve the same DOF in the original fading network as the rate in the polymatroidal network, i.e.,

$$\mathcal{R}_{\text{ach}}^{\text{poly}} = \mathcal{D}_{\text{ach}}^{\text{fading}}. \quad (100)$$

This implies that

$$\mathcal{D}_{\text{ach}}^{\text{fading}} \supseteq \frac{\mathcal{D}_{\text{cut}}^{\text{fading}}}{\mathcal{O}(\log^3 k) \mathcal{O}(\log p_{\max})}. \quad (101)$$

This completes the proof of the theorem. \square

5.6.1 Special Traffic Scenarios

We now present results for *directed* networks with MIMO broadcast and MAC channels with delayed feedback under the special traffic patterns presented in Sec. 3.4. Since the networks are directed, reciprocity is not needed; however we will continue to assume that the broadcast channel gets delayed CSI feedback.

Theorem 19. *For a directed network composed of MIMO broadcast and MAC channels with delayed CSI feedback, a simple separation strategy can achieve a DOF region,*

$$\mathcal{D}_{\text{ach}} \supseteq \frac{\mathcal{D}_{\text{cut}}}{\log(p_{\max} + 1)} \quad \text{for BC Traffic}, \quad (102)$$

$$\mathcal{D}_{\text{ach}}^{\text{sum}} \supseteq \frac{\mathcal{D}_{\text{cut}}^{\text{sum}}}{\log(p_{\max} + 1)} \quad \text{for X Traffic}, \quad (103)$$

$$\mathcal{D}_{\text{ach}}^{\text{sum}} \supseteq \frac{\mathcal{D}_{\text{cut}}^{\text{sum}}}{2 \log(p_{\max} + 1)} \quad \text{for group-communication traffic}, \quad (104)$$

where

$$p_{\max} = \max_{\text{BC channels}} p(\text{BC channel}), \quad (105)$$

and $p(\text{BC channel})$ is given by the minimum of number of transmit antennas and the total number of received antennas in the broadcast channel.

Proof. The proofs are similar to the proof of Theorem 18 and are therefore omitted. \square

5.7 Fading Linear Deterministic Network

Consider a linear deterministic network of the form defined in Sec. 4.7, where each of the non-zero channel coefficients $H_{ij}(t)$ undergo i.i.d. fading with a uniform distribution on the non-zero elements. Then the local scheme of Sec. 4.7 can be extended to a global network scheme.

Theorem 20. *For a bidirected linear deterministic network with k source destination pairs, the rate region given by*

$$\mathcal{R}_{\text{ach}} \supseteq \frac{\mathcal{R}_{\text{cut}}}{\mathcal{O}(\log k)}, \quad (106)$$

is achievable using a separation strategy.

Proof. The proof proceeds analogously to Theorem 9, but instead of using the local scheme in Theorem 2, the scheme of Theorem 4. \square

5.7.1 Special Traffic Scenarios

We now present results for *directed* fast fading linear deterministic networks under the special traffic patterns presented in Sec. 3.4. Since the network is directed, reciprocity will not be necessary to prove this result.

Theorem 21. *For a directed fast fading linear deterministic network, a simple separation strategy can achieve a rate,*

$$\mathcal{R}_{ach} \supseteq \frac{\mathcal{R}_{cut}}{2} \text{ for BC Traffic,} \quad (107)$$

$$R_{ach}^{sum} \supseteq \frac{R_{cut}^{sum}}{2} \text{ for X Traffic,} \quad (108)$$

$$R_{ach}^{sum} \supseteq \frac{R_{cut}^{sum}}{4} \text{ for group-communication traffic.} \quad (109)$$

Acknowledgements

The authors would like to thank Chandra Chekuri and Adnan Raja for several useful discussions.

A Proof of Lemma 4

The rate region for cut-set under product distribution is given by:

$$\mathcal{R}_{\text{cut,product}}^{\text{MAC}}(P) = \left\{ R : \sum_{i \in S} R_i \leq \log(1 + \sum_{i \in S} |h_i|^2 P) \right\}. \quad (110)$$

The rate region for cut-set under general distribution is given by:

$$\mathcal{R}_{\text{cut,general}}^{\text{MAC}}(P) = \left\{ R : \sum_{i \in S} R_i \leq \log(1 + (\sum_{i \in S} |h_i|)^2 P) \right\}. \quad (111)$$

By the Cauchy-Schwarz inequality, we get

$$\left(\sum_{i \in S} |h_i| \right)^2 P \leq \left(\sum_{i \in S} |h_i|^2 \right) dP, \quad (112)$$

which in turn implies

$$\mathcal{R}_{\text{cut,general}}^{\text{MAC}}(P) \subseteq \mathcal{R}_{\text{cut,product}}^{\text{MAC}}(dP). \quad (113)$$

We can similarly show that,

$$\mathcal{R}_{\text{cut,general}}^{\text{BC}}(P) = \mathcal{R}_{\text{cut,product}}^{\text{MAC}}(dP). \quad (114)$$

Along with the equality in (17), this implies that

$$\mathcal{R}_{\text{cut,general}}^{\text{MAC}}(P) \subseteq \mathcal{R}_{\text{ach}}^{\text{MAC}}(dP). \quad (115)$$

We similarly get, using (21) and (114),

$$\mathcal{R}_{\text{cut,general}}^{\text{BC}}(P) \subseteq \mathcal{R}_{\text{ach}}^{\text{BC}}(dP). \quad (116)$$

B Proof of Lemma 5

Without feedback the capacity of this erasure broadcast channel can be easily found. This is because this erasure broadcast channel is stochastically degraded [17], and the capacity is given by

$$\left\{ (R_1, \dots, R_d) \mid \sum_{i=1,2,\dots,d} R_i \leq 1 - \epsilon \right\}. \quad (117)$$

The rate region can be achieved by time sharing between the individual links. We can compare this rate to the cut-set bound which is given by

$$R_{\text{cut}} = \left\{ (R_1, \dots, R_D) \mid \sum_{i \in J} R_i \leq 1 - \epsilon^{|J|} \quad \forall J \subseteq \{1, 2, \dots, d\} \right\}. \quad (118)$$

The ratio between the sum rate of the scheme and the cut-set bound is the factor

$$\frac{1 - \epsilon}{1 - \epsilon^d}, \text{ which } \rightarrow \frac{1}{d}, \text{ as } \epsilon \rightarrow 1. \quad (119)$$

As expected, the time sharing region does not compare very favorably to the cut-set bound.

C Proof of Lemma 7

Consider the rate region with feedback $\mathcal{R}_{\text{ach,fb}}$. We would like to know for what value of A does $A\mathcal{R}_{\text{ach,fb}} \supseteq \mathcal{R}_{\text{cut}}$. Let us take a point in \mathcal{R}_{cut} , we would like to know, for what value of A does this imply $\sum_{i=1,2,\dots,d} \frac{R_i}{1-\epsilon^i} \leq A$. This is equivalent to

$$A = \max \sum_{i=1}^n \frac{R_i}{1-\epsilon^i}, \quad (120)$$

such that,

$$\sum_{i \in J} R_i \leq 1 - \epsilon^{|J|} \quad \forall J \subseteq \{1, 2, \dots, d\}. \quad (121)$$

This is a linear optimization over a polymatroid and the optimal solution is given by the greedy algorithm [22],

$$(R_1, \dots, R_d) = (1 - \epsilon, \epsilon - \epsilon^2, \dots, \epsilon^{n-1} - \epsilon^n), \quad (122)$$

and the optimal value of the objective function is

$$\sum_{i=1}^n \frac{\epsilon^{i-1} - \epsilon^i}{1 - \epsilon^i}. \quad (123)$$

Lets examine the i -th term in this sum, substituting $x = \epsilon^{-1}$,

$$\frac{x-1}{x^i-1} = \frac{1}{1+x+x^2+\dots+x^{i-1}} \leq \frac{1}{i}. \quad (124)$$

Therefore the sum is upper bounded by

$$A \leq \sum_i \frac{1}{i} \leq \log d. \quad (125)$$

D Feedback: Multiple Access Erasure Channel

Consider a finite field multiple access erasure channel, where

$$y = \sum_{i=1}^d e_i x_i, \quad (126)$$

where e_i are i.i.d. Bernouli with probability $1 - \epsilon$. In this channel, some of the transmitters' packets can get erased, and the received vector is the sum of those packets that did not get erased.

This multiple access channel is the dual of the broadcast erasure channel in the sense that the cut-set bound of the two channels are identical. This channel can be realized physically by having using computation code on the wireless channel, which computes the required linear combination. If all the channel coefficients in the wireless channel are good, then the combination $\sum_{i=1}^d X_i$ can be computed easily. However if one of the channel coefficient, say h_j , is in deep fade, then it may be algorithmically hard to compute this linear combination. Therefore, one way around this problem is to avoid having X_j in the linear combination and instead compute $\sum_{i=1,2,\dots,d} \text{ }_{i \neq j} X_i$. This gives rise to the channel model in (126).

The capacity of this multiple access channel is given by the cut-set bound (which is the same as the broadcast channel cut-set bound),

$$\mathcal{R}^{\text{MAC}} = \mathcal{R}_{\text{cut}}^{\text{MAC}} = \{(R_1, \dots, R_D) \mid \sum_{i \in J} R_i \leq 1 - \epsilon^{|J|} \quad \forall J\}. \quad (127)$$

Here $J \subseteq \{1, 2, \dots, d\}$. We emphasize that the capacity region of this multiple access channel is equal to the cut-set bound of the erasure broadcast channel.

E Proof of Lemma 9

We will first consider the case of a network with l transmit antennas and m single-antenna receivers. We can assume $l \leq m$, since if $l > m$, we can restrict ourselves to using m transmit antennas, which leaves the cut unaltered. Therefore $p = l$.

We can choose any particular subset of l receivers and use the strategy in Lemma 8 to achieve a DOF of $\frac{1}{\mathcal{O}(\log l)}$ for each receiver. We can time share between all possible subsets of size l to achieve a certain DOF region. To compute the rate region achievable by this method, we use the following trick: Let the DOF tuple achieved be $\frac{1}{\mathcal{O}(\log l)}(r_1, \dots, r_m)$. Let us construct a bi-partite graph with l nodes on the left partition and m nodes on the right partition and a complete graph connecting them. Each matching is equivalent to choosing a certain subset of the receivers (given by the set of right-partition nodes covered by the matching) and achieving DOF 1 for the each of the receivers. The characteristic vector of a bipartite matching is given by $(x_{ij})_M$ such that $x_{ij} = 1$ for edge (i, j) in the matching M and $x_{ij} = 0$ otherwise. The convex hull of these characteristic vectors is given by

$$\mathcal{M} = \text{conv} \{(x_{ij})_M \mid M \text{ a matching}\}. \quad (128)$$

For a bipartite graph, this is equivalent to the following polytope [77]

$$\mathcal{P} = \left\{ (x_{ij}) \mid x_{ij} \geq 0 \quad \forall i, j, \sum_j x_{ij} \leq 1 \quad \forall i, \sum_i x_{ij} \leq 1 \quad \forall j \right\}. \quad (129)$$

The DOF d_j is given by

$$d_j = \frac{1}{\mathcal{O}(\log l)} \sum_i x_{ij}. \quad (130)$$

Now consider the following polytope,

$$\mathcal{D}_{\text{ach}} = \left\{ (d_j) \mid d_j \geq 0 \quad \forall j, d_j \leq \frac{1}{\mathcal{O}(\log l)}, \sum_j d_j \leq \frac{l}{\mathcal{O}(\log l)} \right\}. \quad (131)$$

We can show that \mathcal{D}_{ach} is equivalent to \mathcal{P} by using the mapping

$$\psi : \mathcal{D}_{\text{ach}} \rightarrow \mathcal{P} \quad (132)$$

$$\psi\{(d_j)\} = (x_{ij}) : x_{ij} = \mathcal{O}(\log l) \frac{d_j}{l}, \quad (133)$$

and the mapping

$$\zeta : \mathcal{P} \rightarrow \mathcal{D}_{\text{ach}} \quad (134)$$

$$\zeta\{(x_{ij})\} = (d_j) : d_j = \frac{1}{\mathcal{O}(\log l)} \sum_i x_{ij}. \quad (135)$$

Thus the region \mathcal{D}_{ach} is achievable. Also the cut-set bound is given by

$$\mathcal{D}_{\text{cut}} = \left\{ (d_j) \mid d_j \geq 0 \quad \forall j, d_j \leq 1, \sum_j d_j \leq l \right\}. \quad (136)$$

This implies that

$$\mathcal{D}_{\text{ach}} = \frac{\mathcal{D}_{\text{cut}}}{\mathcal{O}(\log l)}, \quad (137)$$

which completes the proof of the single antenna receiver case.

For the multi-antenna receiver case, we will treat it as being composed of many single antenna receivers each of which are receiving independent information and then sum up the rates. This proof will extend to this case to get the desired result.

F Proof of Lemma 14

Let us start with a cut on the polymatroidal network. A cut on the polymatroidal network specifies a vertex partition Ω and along with it, specifies how to compute the value of the cut, by specifying which edges to group together for the sub-modular constraints. For example, consider the cut in Fig. 5b, it features the vertex partition and also specifies how to group edges to get an upper bound.

From the given partition we need to construct a cut on the Gaussian network. The polymatroidal cuts specify which of the edges need to be grouped together. By bounding the polymatroidal network in this manner, each edge is involved in *either* a broadcast constraint or a superposition constraint. This is in contrast to the Gaussian case, where each cut has a certain value, and there is no sense in which edges are assigned to broadcast or superposition constraint.

The key idea to connect these two cuts is the idea of *decoupling* the constraints in the Gaussian channel:

- In the Gaussian channel, if broadcast constraint is not active for a given edge, then the edge only participates in the superposition constraint and vice versa.
- While this is not true as it is in the Gaussian network, we can obtain an upper bound network where certain broadcast or superposition constraints are no longer active. This process of obtaining upper bound network where a certain constraint is not active is called decoupling.
- Decoupling a *broadcast constraint* is easy, because a network in which edges are not involved in a broadcast constraint can only, in general, do better than a network where there is a broadcast constraint on the edges.
- Decoupling the *superposition constraint* requires a bit more work, this can be illustrated using the following example. Suppose two edges e_1 and e_2 are involved in the superposition constraint in the following manner:

$$y = x_1 + x_2 + z. \quad (138)$$

Then if we have construct another channel in which the edges e_1 and e_2 are not involved in a broadcast constraint, then we have,

$$y_i = x_i + z_i, i = 1, 2, \quad (139)$$

this channel can emulate any scheme in the original channel if $\text{Var}(z_i) = \frac{\text{Var}(z)}{2}$. If this condition is satisfied, then we can add up $y_1 + y_2$ to get $x_1 + x_2 + (z_1 + z_2)$, which is statistically equivalent to the original channel. Therefore to decouple the superposition constraint involving d variables, we need to reduce the variance of the noise by d , the degree of the superposition constraint, or equivalently, increase the signal power by a factor d .

- Therefore we can decouple all required broadcast and superposition constraints, if the power is increased by a factor of d , which is the maximum degree of any node.

Thus given any cut in the polymatroidal network along with the assignments of the edges to broadcast or superposition constraints, we can obtain a similar cut in the Gaussian network by decoupling the constraints which are not active in the polymatroidal cut. This incurs a power penalty factor of d , thus as far as the outer bound is concerned, we can assume that each node has power dP instead of P . The network thus obtained is made of MAC and broadcast channels. In this network, every cut decomposes into the sum of MAC and BC cuts. A MAC cut with d nodes when evaluated under general distribution on the input, is of the form,

$$\sum_i R_{ij} \leq \mathbb{E} \log \left(1 + \left(\sum_i |h_{ij}| \right)^2 dP \right) \quad (140)$$

$$\leq \mathbb{E} \log \left(1 + \sum_i |h_{ij}|^2 d^2 P \right) \quad \text{by Cauchy-Schwartz inequality} \quad (141)$$

$$\leq \log \left(1 + \mathbb{E} \left(\sum_i |h_{ij}|^2 \right) d^2 P \right) \quad \text{by Jensen} \quad (142)$$

$$= \log (1 + d^3 P) \quad (143)$$

$$\leq \mathbb{E} \{ \log (1 + ad^3 P |h|^2) \} = C(ad^3 P) \quad (144)$$

The last step follows because,

$$\mathbb{E} \{ \log (1 + c|h|^2) \} = \mathbb{E} \left\{ \log \left(1 + ce^{\log |h|^2} \right) \right\} \quad (145)$$

$$\geq \log \left(1 + ce^{\mathbb{E}(\log |h|^2)} \right). \quad (146)$$

Here $a := e^{-\mathbb{E}(\log |h|^2)}$ is finite for the fading distribution, by assumption (it is equal to 1). Thus, the cutset bound for the original network implies that

$$\sum_i R_{ij} \leq C(ad^3 P), \quad (147)$$

whereas the corresponding cutset bound for the polymatroidal network is of the form,

$$\sum_i R_{ij} \leq \frac{1}{2} C(2P). \quad (148)$$

We get,

$$\mathcal{R}_{\text{cut}}^{\text{poly}} \supseteq \frac{1}{2} \mathcal{R}_{\text{cut}}^{\text{original}} \left(\frac{P}{bd^3} \right), \quad (149)$$

where $b := \frac{a}{2} = \frac{e^{\mathbb{E}(\log |h|^2)}}{2}$ is a constant depending on the fading distribution. For h distributed as complex gaussian, $b \approx 0.86$.

G Proof of Theorem 13

The k sources are in the layer V_0 , and the k destinations are in the layer V_{L+1} . Let the number of nodes in the i -th layer be n_i .

G.0.2 Max-Flow Rate

Let R_i be the rate between the i -th source destination pair. We will route the flow in a symmetric manner, where the incoming flow is divided equally among all the edges going out of a node. We will compute constraints on the rate region achievable by this strategy.

In the first hop, all edges going out of the i -th source will carry a flow of value $\frac{R_i}{n_1}$. The constraint imposed by the edges going out of the source is given by

$$\frac{R_i}{n_1} n_1 \leq 1 \iff R_i \leq 1, \quad \forall i = 1, 2, \dots, k. \quad (150)$$

The constraint imposed by the edges coming into the nodes of the first hop are given by

$$\frac{\sum_i R_i}{k n_1} k \leq 1 \iff \sum_i R_i \leq n_1. \quad (151)$$

The total flow carried by all the nodes in any given layer equals $\sum_i R_i$, and each node in layer l carries a flow of $\frac{R_i}{n_l}$ corresponding to flow i .

In the l -th hop connecting layers V_{l-1} and V_l , each edge carries a flow of value $\frac{R_i}{n_l n_{l+1}}$ corresponding to source i , which yields a total flow of value $\frac{\sum_i R_i}{n_l n_{l+1}}$ for each edge. The outgoing constraints on layer l yield,

$$\frac{\sum_i R_i}{n_l n_{l+1}} n_{l+1} \leq 1 \iff \sum_i R_i \leq n_l, \quad \forall i = 1, 2, \dots, k. \quad (152)$$

The incoming constraints on layer l yield,

$$\frac{\sum_i R_i}{n_l n_{l+1}} n_l \leq 1 \iff \sum_i R_i \leq n_{l+1}, \quad \forall i = 1, 2, \dots, k. \quad (153)$$

In the final $(L + 1)$ -th hop also, there are constraints similar to layer 1. In particular, the outgoing constraints on layer l yield,

$$\frac{\sum_i R_i}{n_L k} k \leq 1 \iff \sum_i R_i \leq n_L. \quad (154)$$

and the incoming constraints on the destination layer yield,

$$\frac{R_i}{n_L} n_L \leq 1 \iff R_i \leq 1, \quad \forall i = 1, 2, \dots, k. \quad (155)$$

Thus a rate pair (R_1, \dots, R_k) is achievable by routing iff

$$\sum_i R_i \leq \min(n_l, n_{l+1}) \quad \forall l = 0, 1, \dots, L. \quad (156)$$

$$R_i \leq 1 \quad \forall i = 1, 2, \dots, k. \quad (157)$$

G.0.3 Cut-set Region

We can easily write the following constraints, which are a subset of the cut-set bounds.

Corresponding to the cut separating $\Omega = \cup_{i=0}^l V_i$ and Ω^c , the following constraint can be written,

$$\sum_i R_i \leq \min(n_l, n_{l+1}) \quad \forall l = 0, 1, \dots, L. \quad (158)$$

Corresponding to the cut given by $\Omega = s_i$, we can get the following constraint

$$R_i \leq 1 \quad \forall i = 1, \dots, k. \quad (159)$$

Comparing (156), (157) and (158), (159), we can deduce that any rate tuple that satisfies the cut-set region will lie in the rate region achieved by the flow. Thus the rate region corresponding to max-flow equals the rate region corresponding to cut-set region.

For example, Fig. 6b denotes the directed layered polymatroidal network obtained from the network in Fig. 6a. Every node basically constrains the total inflow and outflow to be lesser than 1.

References

- [1] C. Chekuri, S. Kannan, A. Raja and P. Viswanath. Multicommodity Flows and Cuts in Polymatroidal Networks. *Proc. Innovations in Theoretical Computer Science*, 2012, to appear. Available online: <http://arxiv.org/abs/1110.6832>
- [2] S. Kannan, A. Raja and P. Viswanath. Approximately Optimal Broadcasting-cum-Multicasting in Wireless Networks. *Proc. of IEEE ISIT*, Aug. 2011.
- [3] S. Kannan, A. Raja and P. Viswanath. Approximately Optimal Broadcasting in Wireless Networks. *IEEE Trans. Inform. Theory*, submitted.
- [4] S. Kannan, A. Raja and P. Viswanath. Local Phy + Global Flow: A Layering Principle for Wireless Networks. *Proc. of IEEE ISIT*, Aug. 2011.
- [5] S. Kannan and P. Viswanath. Multiple-Unicast in Fading Wireless Networks: A Separation Scheme is Approximately Optimal. *Proc. of IEEE ISIT*, Aug. 2011.
- [6] R. Ahlswede, C. Ning, S. Li, and R. Yeung. Network information flow. *IEEE Trans. Inform. Theory*, vol. 46, no. 4, pp. 1204-1216, July 2000.
- [7] S. Aeron and V. Saligrama. Wireless ad hoc networks: Strategies and scaling laws for the xed SNR regime. *IEEE Transactions on Information Theory*, vol. 53, no. 6, pp. 2044–2059, Jun. 2007.
- [8] M. A. M. Ali and D. Tse. Completely Stale Transmitter Channel State Information is Still Very Useful. *Proc. 48th Annual Allerton Conference*, Oct. 2010.
- [9] A. Amadruz, and C. Fragouli. Combinatorial algorithms for wireless information flow. *Proc. of ACM-SIAM SODA*, Jan. 2009.
- [10] M. Anand and P. R. Kumar, A Digital Interface for Gaussian Relay and Interference Networks: Lifting Codes From the Discrete Superposition Model. *IEEE Trans. Inform. Theory*, vol. 57, no. 5, pp. 2548 - 2564.

- [11] V. S. Annapureddy, A. El Gamal, and V. V. Veeravalli. Degrees of Freedom of Interference Channels with CoMP Transmission and Reception. *Available Online*, <http://arxiv.org/abs/1109.5396>
- [12] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse. Wireless Network Information Flow: A Deterministic Approach. *IEEE Trans. Info. Theory.*, 57(4):1872–1905, 2011.
- [13] J. Bourgain. On Lipschitz embedding of finite metric spaces in Hilbert space. *Israeli J. Math.*, 52:46–52, 1985.
- [14] V. Cadambe and S. Jafar. Interference alignment and the degrees of freedom of the K user interference channel *IEEE Trans. Inform. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.
- [15] T. H. Chan and A. Grant. Mission impossible: Computing the network coding capacity region. *Proc. IEEE Int. Symp. Inf. Theory*, Toronto, ON, Canada, Jul. 2008
- [16] J. Chuzhoy and S. Khanna. Polynomial Flow-Cut Gaps and Hardness of Directed Cut Problems. *Journal of the ACM*, Volume 56, issue 2, Article 6, 2009. Preliminary version in *Proc. of the ACM STOC*, 2007.
- [17] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd Edition, John Wiley and Sons, New York, 2006.
- [18] A. F. Dana, M. Sharif, R. Gowaikar, B. Hassibi, and M. Effros, Is broadcast plus multiaccess optimal for Gaussian wireless networks? *Asilomar Conf. on Signals, Systems and Computers*, 2003.
- [19] A. F. Dana, R. Gowaikar, R. Palanki, B. Hassibi, and M. Effros. Capacity of wireless erasure networks. *IEEE Trans. Inform. Theory*, vol. 52, no. 3, pp. 789–804, March 2006.
- [20] C. Tian, J. Chen, S. N. Diggavi, and S. Shamai, Optimality and Approximate Optimality of Source-Channel Separation in Networks. *Proc. IEEE ISIT*, pp. 495–499, Austin, June, 2010.
- [21] R. Etkin, D. Tse, and H. Wang, “Gaussian Interference Channel Capacity to Within One Bit,” *IEEE Trans. Inform. Theory*, vol. 54, no. 12, pp. 5534–5562, Dec. 2008.
- [22] J. Edmonds, “Submodular functions, matroids and certain polyhedra, *Proc. Calgary Int. Conf. Combinatorial Structures and Applications*. Calgary, Alta, June 1969, pp. 69–87.
- [23] J. Edmonds and E. Giles A min-max relation for submodular functions on graphs. *Annals of Discrete Mathematics*, vol.1, 185–204, 1977.
- [24] R. Etkin and E. Ordentlich. On the degrees-of-freedom of the K-user Gaussian interference channel. *Available Online*, <http://arxiv.org/abs/0901.1695>, Jan 2009.
- [25] A. Federgruen and H. Groenevelt, Polymatroidal flow network models with multiple sinks. *Networks*, 18(4): 285–302, 1988.
- [26] U. Feige, M.T. Hajiaghayi, and J. Lee. Improved approximation algorithms for minimum-weight vertex separators. *SIAM J. on Computing*, 38(2): 629–657, 2008. Preliminary version in *Proc. of ACM STOC*, 563–572, 2005.
- [27] M. Franceschetti, M. D. Migliore, P. Minero, and F. Schettino. The Degrees of Freedom of Wireless Networks Via Cut-Set Integrals. *IEEE Trans. Inform. Theory*, vol. 57, no. 5, pp. 3067 – 3079.
- [28] A. El Gamal, On Information Flow in Relay Networks, *IEEE National Telecomm. Conf.*, Vol. 2, pp. D4.1.1 - D4.1.4, Nov. 1981.

- [29] M. Gatzianas, L. Georgiadis, and L. Tassiulas. Multiuser broadcast erasure channel with feedback - capacity and algorithms. *Available Online*, <http://arxiv.org/abs/1009.1254>, 2010
- [30] L. Georgiadis and L. Tassiulas. Broadcast erasure channel with feedback capacity and algorithms. *Proc. 5th Workshop on Net. Cod. Theory and Appl.*, June 2009, pp. 54-61.
- [31] M. X. Goemans, S. Iwata, and R. Zenklusen. A flow model based on polylinking system. *Math. Programming Ser A*, to appear. Preliminary version in *Proc. Allerton Conf*, Sep. 2009
- [32] T. Gou, S. A. Jafar, S.-W. Jeon, and S.-Y. Chung. Aligned Interference Neutralization and the Degrees of Freedom of the $2 \times 2 \times 2$ Interference Channel. *Available Online*, <http://arxiv.org/abs/1012.2350v1>
- [33] P. Gupta and P. R. Kumar, Capacity of Wireless Networks *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 388-404, March 2000.
- [34] D. Gunduz, A. Yener, A. Goldsmith, and H. V. Poor. The Multi-way Relay Channel. *Available Online*, <http://arxiv.org/abs/1004.2434>
- [35] N. J. A. Harvey, R. Kleinberg, and A. R. Lehman, "On the capacity of information networks," *IEEE Trans. Inf. Theory*, vol. 52, pp. 2345-2364, June 2006.
- [36] R. Hassin. On Network Flows. Ph.D Dissertation, Yale University, 1978.
- [37] Syed A. Jafar. Interference Alignment: A New Look at Signal Dimensions in a Communication Network. *Foundations and Trends in Communications and Information Theory*, Vol. 7, No. 1, pages: 1-136.
- [38] S. -W. Jeon, S. -Y. Chung, and S. Jafar, "Degrees of freedom of multi-source wireless relay networks," *Proc. Allerton Conf.*, Sep. 2009.
- [39] S. Jafar and S. Shamai. Degrees of Freedom Region of the MIMO X Channel. *IEEE Trans. Inform. Theory*, vol. 54, no. 1, pp. 151-170, Jan. 2008.
- [40] S. -W. Jeon; S. -Y. Chung; S. A. Jafar. Degrees of Freedom Region of a Class of Multisource Gaussian Relay Networks. *IEEE Trans. Inform. Theory*, vol. 57, no. 5, pp. 3032 – 3044.
- [41] S. Kamath, V. Ananthram and D. Tse. Generalized Network Sharing Outer Bound and the Two-Unicast Problem. *Proc. NetCod*, July 2011.
- [42] M. Kim, E. Erez, E. M. Yeh, and M. Mdard. Deterministic Network Model Revisited: An Algebraic Network Coding Approach. *Available Online*, <http://arxiv.org/abs/1103.0999>
- [43] G. Kramer and S. A. Savari. Edge-cut bounds on network coding rates. *J. Network and Systems Management*, vol. 14, no. 1, pp. 49-67, March 2006.
- [44] R. Koetter and M. Mdard. An algebraic approach to network coding. *IEEE Trans. Networking*, vol.11, no.5, Oct. 2003.
- [45] R. Koetter, M. Effros, and M. Mdard. A Theory of Network Equivalence – Part I: Point to Point Channels. *IEEE Trans. Inform. Theory*, Feb. 2011.
- [46] R. Koetter, M. Effros, and M. Mdard. A Theory of Network Equivalence – Part II: Multiterminal Channels. *Available Online*, <http://arxiv.org/abs/1007.1033>

- [47] W.-C. Kua and C.-C. Wang. On the Capacity of 2-User 1-Hop Relay Erasure Networks the Union of Feedback, Scheduling, Opportunistic Routing, and Network Coding. *Proc. IEEE ISIT*, Aug. 2011.
- [48] M. Langberg and M. Mardar, The Multiple Unicast Network Coding Conjecture. *Proc. Allerton Conf. Commun., Control, and Computing*, Monticello, IL, October 2009.
- [49] E. L. Lawler and C. U. Martel. Computing maximal “Polymatroidal” network flows. *Math. Oper. Res.*, 7(3):334–347, 1982.
- [50] F. T. Leighton and S. Rao. Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms. *JACM*, 46(6):787–832, 1999. Preliminary version in *Proc. of ACM STOC*, 1988.
- [51] S. H. Lim; Y. -H. Kim; A.. El Gamal, and S. -Y. Chung. Noisy Network Coding. *IEEE Trans. Inform. Theory*, vol. 57, no. 5, pp.3132–3152.
- [52] Z. Li and B. Li. Network Coding: The Case of Multiple Unicast Sessions *Proc. 42nd Allerton Conference*, Monticello, IL, 2004.
- [53] D. S. Lun, M. Mardar, R. Koetter, and M. Effros, On coding for reliable communication over packet networks, *Physical Communication*, vol. 1, no. 1, March 2008, pp. 3–20.
- [54] S. Shakkottai and R. Srikant. Network Optimization and Control. *Foundations and Trends in Networking*, NoW Publishers, 2007.
- [55] M. Chiang, S. H. Low, A. R. Calderbank, and J. C. Doyle. Layering as Optimization Decomposition: A Mathematical Theory of Network Architectures. *Proceedings of the IEEE*, pp.255–312, vol. 95, no. 1, Jan. 2007.
- [56] O. Leveque and I. E. Telatar. Information-theoretic upper bounds on the capacity of large, extended ad hoc wireless networks. *IEEE Transactions on Information Theory*, vol. 51, no. 3, pp. 858–865, Mar. 2005.
- [57] N. Linial, E. London, and Y. Rabinovich. The Geometry of Graphs and Some of its Algorithmic Applications. *Combinatorica*, 15(2):215–245, 1995.
- [58] L. Lovász. Submodular functions and convexity. *Mathematical programming: the state of the art*, pages 235–257, 1983.
- [59] M. A. Maddah-Ali, S. A. Motahari, and Amir K. Khandani. Communication over MIMO X Channels: Signaling and Performance Analysis. *IEEE Trans. on Inform. Theory*, vol. 54, no. 8, pp. 3457–3470, 2008.
- [60] K. Marton, “A coding theorem for the discrete memoryless broadcast channel,” *IEEE Trans. Inform. Theory*, vol. 25, no. 3, pp. 306–311, 1979.
- [61] J. Naor and L. Zosin. A 2-approximation algorithm for the directed multiway cut problem. *SIAM J. on Computing*, 31(2):477–482, 2001. Preliminary version in *Proc. of IEEE FOCS*, 1997.
- [62] J. Matousek and Y. Rabinovich. On dominated ℓ_1 metric. *Israeli J. of Mathematics*, 123:285–301, 2001.
- [63] S. Mohajer, S. N. Diggavi, C. Fragouli, and D. N. C. Tse. Approximate Capacity of Gaussian Interference-Relay Networks. *IEEE Trans. Inform. Theory*, vol. 57, no. 5, pp. 2837–2684.

- [64] A. Motahari, S. O. Gharan, and A. Khandani. Real interference alignment with real numbers. *Available Online*, <http://arxiv.org/abs/0908.1208>, Aug. 2009.
- [65] B. Nazer, M. Gastpar, S. Jafar, and S. Vishwanath. Ergodic interference alignment. *Proc. of Intl. Symp. Inform. Theory*, Seoul, 2009.
- [66] D O'Neill, B.S. Thian, A. Goldsmith and S. Boyd. Wireless NUM: Rate and Reliability Tradeoffs in Random Environments *Proc. IEEE WCNC*, Budapest Hungary, 2009.
- [67] U. Niesen “Interference Alignment in Dense Wireless Networks.” *Available Online*, <http://arxiv.org/abs/0912.0868>
- [68] U. Niesen, P. Gupta, D. Shah. The Balanced Unicast and Multicast Capacity Regions of Large Wireless Networks. *IEEE Transactions on Information Theory*, May 2010.
- [69] U. Niesen, P. Gupta, D. Shah. On Capacity Scaling in Arbitrary Wireless Networks. *IEEE Transactions on Information Theory*, September 2009.
- [70] U. Niesen, B. Nazer and P. Whitting. Computation Alignment: Capacity Approximation without Noise Accumulation. *Available Online*, <http://arxiv.org/abs/1108.6312>
- [71] J. Oxley. *Matroid Theory*, Oxford University Press, 2006.
- [72] O. Oyman, R. U. Nabar, H. Bolcskei, and A. J. Paulraj, “Tight lower bounds on the ergodic capacity of Rayleigh fading MIMO channels,” *Proc. IEEE Globecom*, vol. 2 , pp.1172 – 1176 , 2002.
- [73] A. Ozgur, O. Leveque and D. N. C. Tse. Hierarchical Cooperation Achieves Optimal Capacity Scaling in Ad Hoc Networks. *IEEE Trans. Inf. Theory*, vol 53, no. 10, pp. 3549 - 3572, October 2007.
- [74] A. Ozgur, R. Johari, D. Tse, and O. Leveque, Information Theoretic Operating Regimes of Large Wireless Networks. *IEEE Transactions on Information Theory*, 56(1):427-437, 2010.
- [75] Y. Rabinovich. On Average Distortion of Embedding Metrics into the Line. *Discrete & Computational Geometry*, 39(4): 720–733, 2008. Preliminary version in *Proc. of ACM STOC*, 2003.
- [76] A. Raja and P. Viswanath. Compress-and-Forward Scheme for a Relay Network: Approximate Optimality and Connection to Algebraic Flows *Proc. of IEEE ISIT*, Aug. 2011.
- [77] A. Schrijver. Combinatorial Optimization: Polyhedra and Efficiency. Springer-Verlag, 2003.
- [78] A. Schrijver. Matroids and linking systems. *Journal of Combinatorial Theory B*, 26:349–369, 1979.
- [79] C. E. Shannon. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* vol. 27, pp. 379–423.
- [80] I. Shomorony and A. S. Avestimehr. Two-Unicast Wireless Networks: Characterizing the Degrees-of-Freedom. *Available Online*, <http://arxiv.org/abs/1102.2498>
- [81] C. Suh and D. N. Tse. Feedback Capacity of the Gaussian Interference Channel to Within 2 Bits. *IEEE Trans. Inform. Theory*, vol. 57, no. 5, pp.2667–2685.
- [82] D. Tse and S. Hanly. Multi-Access Fading Channels: Part I: Polymatroid Structure, Optimal Resource Allocation and Throughput Capacities. *IEEE Transactions on Information Theory*, vol. 44, No. 7, pp. 2796–2815, November 1997.

- [83] D. Tse and P. Viswanath, “*Fundamentals of Wireless Communication*,” Cambridge University Press, New York, 2005.
- [84] A. Vahid, V. Aggarwal, A. S. Avestimehr, and A. Sabharwal. On the Capacity of Multi-Hop Wireless Networks with Partial Network Knowledge. *Available Online*, <http://arxiv.org/abs/1010.1060>
- [85] D. Vasudevan and S. Korada. Polymatroidal Flows on Two Classes of Information Networks. *IEEE Transactions on Information Theory*, vol. 57, no.1, pp. 227–233, 2011.
- [86] C. S. Vaze and M. K. Varanasi. The Degrees of Freedom of the 2-Hop, 2-User Interference Channel with Feedback. *Proc. Allerton Conf.*, Sep. 2011.
- [87] P. Viswanath and D. Tse, “Sum capacity of the multiple antenna Gaussian broadcast channel and uplink-downlink duality,” *IEEE Trans. Inf. Theory*. vol. 49, no. 8, pp. 1912–1923, Aug. 2003.
- [88] S. Vishwanath, N. Jindal, and A. J. Goldsmith, “Duality, achievable rates and sum-rate capacity of Gaussian MIMO broadcast channel,” *IEEE Trans. Inf. Theory*. vol. 49, pp. 2658–2668, 2003.
- [89] C.-C. Wang. Capacity of 1-to-K broadcast packet erasure channels with channel output feedback. *Proc. 48th Annual Allerton Conference*, Oct. 2010.
- [90] I. Wang and D. N. Tse. Interference Mitigation Through Limited Receiver Cooperation. *IEEE Trans. Inform. Theory*, vol. 57, no. 5, pp. 2913 - 2940 .
- [91] I. Wang and D. N. Tse. Interference Mitigation Through Limited Transmitter Cooperation. *IEEE Trans. Inform. Theory*, vol. 57, no. 5, pp. 2941 - 2965 .
- [92] H. Weingarten, Y. Steinberg, S. Shamai, “The Capacity Region of the Gaussian MIMO Broadcast Channel”, *IEEE Trans. Inform. Theory*, vol. 52(9), Sept. 2006.
- [93] Y. Wu, S. Shamai and S. Verdú. Degrees of Freedom of the Interference Channel: a General Formula. *Proc. ISIT*, Aug. 2011.
- [94] F. Xue, L.-L. Xie, and P. R. Kumar. The transport capacity of wireless networks over fading channels. *IEEE Transactions on Information Theory*, vol. 51, no.3, pp. 834–847, Mar. 2005.
- [95] L.-L. Xie and P. R. Kumar. On the path-loss attenuation regime for positive cost and linear scaling of transport capacity in wireless networks. *IEEE Transactions Information Theory*, vol. 52, no. 6, pp. 2313–2328, Jun. 2006.
- [96] R. Yates and D. Tse. Fading Broadcast Channels with State Information at the Receivers. *Available Online*: <http://arxiv.org/abs/0904.3165>
- [97] S. M. S. Yazdi and S. A. Savari. A Max-Flow/Min-Cut Algorithm for Linear Deterministic Relay Networks. *IEEE Trans. on Info. Theory*, 57(5):3005–3015, 2011. Preliminary version in *Proc. of ACM-SIAM SODA*, 2010.